# Skriptum zur Vorlesung
# Mathematische Logik

Klaus Gloede
Mathematisches Institut der Universität Heidelberg

Wintersemester 2006/07

# Inhaltsverzeichnis

# Teil I

# Collegium Logicum

Mein teurer Freund, ich rat´ Euch drum
Zuerst Collegium Logicum.
Da wird der Geist Euch wohl dressiert
In spanische Stiefeln eingeschnürt,
Daß er bedächtiger so fortan
Hinschleiche die Gedankenbahn.

*J. W. v. Goethe: Faust I*

Alle Menschen sind sterblich, Sokrates ist ein Mensch,
also ist Sokrates sterblich.

So lautet ein bekanntes Beispiel eines logischen Schlusses. Dabei sagt die Logik nichts aus über die Gültigkeit der Einzelaussagen, insbesondere behauptet sie nicht die Sterblichkeit von Sokrates, sondern die Wahrheit dieser Aussage unter Annahme, daß die beiden vorangehenden Aussagen (die Prämissen) wahr sind. Nach dem gleichen Muster könnte man auch folgern:

Alle Tappel haben einen Deppel, Knobbel ist ein Tappel,
also hat Knobbel einen Deppel.

Bereits ARISTOTELES (384 - 322 v. Chr.) hat in seiner Syllogistik damit begonnen, derartige Schlußweisen zu formalisieren und alle möglichen Schlüsse systematisch auf ihre Gültigkeit zu untersuchen[1]. Auch erste Ansätze einer Aussagenlogik finden sich bei ARISTOTELES ; sie wird vor allem in der Stoa und in der Scholastik weiterentwickelt. Den obigen Schluß können wir in moderner Symbolik in der Form

$$\forall X(M(X) \to T(X)), M(S)$$
$$\text{also: } T(S)$$

[1] I. Susan Russinoff hat 1999 darauf hingewiesen, daß die endgültige Klassifizierung der Syllogismen erst 1883 von einer Schülerin des amerikanischen Logikers C.S. Peirce dargestellt wurde: Christine Ladd-Franklin. Bereits 1882 hatte sie alle Voraussetzungen für eine Promotion an der John Hopkins University erfüllt; da sie aber damals als Frau nicht zugelassen wurde, erhielt sie ihren Ph.D.-Grad erst 1926, und ihre Arbeit blieb über hundert Jahre unbeachtet.

schreiben. Daraus wird deutlich, daß der Schluß ganz allgemein und allein aufgrund seiner Form gilt - die Untersuchung derartiger Schlußweisen ist Gegenstand der Mathematischen Logik (auch Formale Logik, Symbolische Logik oder Metamathematik, von Philosophen auch Moderne Logik oder Logistik genannt).

Die Mathematische Logik ist

(i) eine Logik für die Mathematik; sie untersucht Ausdrucksmöglichkeiten in formalen Sprachen und Schlußweisen, wie sie in der Mathematik insbesondere für axiomatische Theorien gebräuchlich sind. Das bedeutet - thematisch - natürlich eine starke Einschränkung gegenüber der allgemeinen (philosophischen) Logik, aber - methodisch - ermöglicht diese Beschränkung es andererseits, sie selbst als

(ii) eine formale mathematischen Theorie zu entwickeln. Dadurch kann die Mathematische Logik (als Metatheorie) über die Mathematische Logik selbst (und zwar nun als formale mathematische Theorie) Aussagen machen. Diese sind möglicherweise rückbezüglich und damit häufig widerspruchsvoll (z. B.: Antinomie des Lügners); in diesem Fall jedoch kann man sie nach K. GÖDEL benutzen, um die Grenzen ihrer eigenen Methode aufzeigen (etwa in Form der GÖDELschen Unvollständigkeitssätze).

# Kapitel 1

# Die Aussagenlogik

Die Aussagenlogik untersucht die Verknüpfungen "nicht", "und", "oder", "wenn …dann" und "genau dann, wenn". Diese Operationen kommen auch in der Umgangssprache vor, allerdings möglicherweise mit zusätzlichen Aspekten, die über den Gebrauch in der Mathematik hinausgehen und hier natürlich unberücksichtigt bleiben sollen. Aufgrund dieser Einschränkung wird die Aussagenlogik (AL) zu einem einfachen Kalkül (dem *Aussagenkalkül*) der beiden Wahrheitswerte $W$ und $F$; mathematisch gesehen handelt es sich um eine sehr einfache Theorie - die Algebra der zwei Werte 0 und 1. Wir wollen sie hier trotzdem ausführlicher behandeln, denn

- die AL ist Grundlage der Computerlogik mit zahlreichen Anwendungen (etwa auf die Untersuchung von Schaltungen) und
- sie ist Bestandteil fast jeder weiteren Logik, insbesondere der später einzuführenden Prädikatenlogik (PL).
- Grundlegende Begriffsbildungen, die besonders für die PL wichtig werden, können hier in einer einfachen Situation vorbereitet werden.

Beginnen werden wir mit dem Aufbau einer *formalen Sprache*: Aus einer vorgegebenen Liste von Zeichen oder

- *Symbolen* (entsprechend dem Alphabet einer natürlichen Sprache) werden nach bestimmten formalen Gesetzen Zeichenreihen gebildet und als
- *Formeln* (entsprechend den Wörtern einer natürlichen Sprache) ausgezeichnet; die Bedeutung der Formeln wird dann durch
- *Interpretationen* (Modelle) festgelegt.

## 1.1 Syntax der Aussagenlogik

Mit Hilfe der Symbole

$$\neg \text{ (nicht)}, \wedge \text{ (und)}, \vee \text{ (oder)}, \rightarrow \text{ (wenn \dots dann)}, \leftrightarrow \text{ (genau dann \dots wenn)}$$

und den Aussagenvariablen (für aussagenlogisch nicht weiter zerlegbare Grundaussagen) definieren wir den Begriff einer *aussagenlogischen Formel*, abgekürzt a. F., wie folgt:

### 1.1.1 Definition der aussagenlogischen Formeln

(F1) $A_0, A_1, A_2, \dots$ *sind a. F.*, und zwar *Aussagenvariable.*

(F2) *Ist* $\varphi$ *eine a. F., so auch* $\neg\varphi$.

(F3) *Sind* $\varphi$ *und* $\psi$ *a. F., so auch* $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$ und $(\varphi \leftrightarrow \psi)$.

(F4) *Das sind alle a. F.*

Aussagenlogische Formeln werden wir wie oben mit kleinen griechischen Buchstaben $\varphi, \psi, \chi, \dots$ bezeichnen, Aussagenvariable auch mit $A, B, C, \dots$

**Achtung:** Formeln sind *formale endliche Zeichenreihen*; ihre *Bedeutung* haben wir zwar oben bereits in Klammern angegeben, aber davon werden (und dürfen) wir zunächst keinen Gebrauch machen!

Beispiele für aussagenlogische Formeln sind:

$$A,\ \neg A,\ \neg\neg\neg B,\ (((\neg\neg A \vee B) \wedge \neg A) \rightarrow (A \wedge \neg C)),$$

während $A\neg,\ B\vee),\ A \wedge \neg C$ keine aussagenlogischen Formeln sind!

#### Klammerregeln

Klammern dienen der eindeutigen Darstellbarkeit von Formeln, können zur besseren Lesbarkeit jedoch in einigen Fällen eingespart werden:

- Äußere Klammern können fortgelassen werden: Schreibe

$$A \wedge \neg C \ \text{ statt } \ (A \wedge \neg C).$$

- $\wedge$ und $\vee$ "binden stärker" als $\rightarrow$ und $\leftrightarrow$: schreibe

$$B\wedge\neg A\rightarrow A\vee\neg C \text{ statt } ((B\wedge\neg A)\rightarrow(A\vee\neg C)).$$

Dagegen werden wir $\wedge$ und $\vee$ gleichberechtigt behandeln, d. h. bei

$$(\neg\neg A\vee B)\wedge\neg A \text{ sowie } (A\vee B)\wedge C \text{ bzw. } A\vee(B\wedge C).$$

kann man keine weiteren Klammern einsparen.

- Für $\wedge$ und $\vee$ benutzen wir die assoziative Schreibweise:

$$A\vee B\vee C \text{ steht für } (A\vee(B\vee C)) \text{ und } A\wedge B\wedge C \text{ für } (A\wedge(B\wedge C)).$$

Die andere mögliche Klammerung führt formal zu einer anderen, inhaltlich jedoch äquivalenten Formel. Rechtsklammerung werden wir später auch wählen bei

- $$A\rightarrow B\rightarrow C \quad \text{für} \quad A\rightarrow(B\rightarrow C),$$

welches jedoch nicht äquivalent zu $(A\rightarrow B)\rightarrow C$ ist! Gelegentlich findet man weitere Vereinbarungen über das Fortlassen von Klammern; gänzlich entbehrlich sind sie bei der *polnischen Notation*: Hier schreibt man etwa

$$N\varphi \text{ für } \neg\varphi, \quad A\varphi\psi \text{ für } \varphi\vee\psi, \quad K\varphi\psi \text{ für } \varphi\wedge\psi,$$

so daß $\neg(p\vee q)\wedge(p\wedge\neg\neg q)$ in polnischer Notation $KNApqKpNNq$ geschrieben wird.

### 1.1.2 Induktion und Rekursion

Eine *explizite* Definition des Begriffs "a. F." würde diesen auf bereits bekannte zurückführen und von der Form sein

$$\varphi \; ist \, a.F. :\Longleftrightarrow \ldots,$$

wobei rechts eine Eigenschaft von $\varphi$ steht (das Definiens), in welcher das Definiendum, nämlich a. F., selbst nicht vorkommt. Dagegen ist die Definition 1.1.1 *rekursiv*, indem zwar auf der rechten Seite auch das zu Definierende vorkommt, aber zurückgeführt auf kürzere Ausdrücke.

Eine rekursive Definition hat den Vorzug, daß sie in der Regel zugleich ein Verfahren enthält, in endlich vielen Schritten "effektiv" nachzuprüfen, ob der vorgelegte Ausdruck der Definitionsanforderung genügt oder nicht. Man sagt in diesem Fall auch: Das Prädikat

$$z_1 \dots z_n \text{ ist eine aussagenlogische Formel}$$

ist *entscheidbar*, d. h. es gibt ein “mechanisches” Verfahren, das - angewandt auf eine vorgegebene endliche Zeichenreihe $z_1 \dots z_n$ - nach endlich-vielen Schritten feststellt, ob $z_1 \dots z_n$ eine aussagenlogische Formel ist oder nicht. Eine Präzisierung der Begriffe “entscheidbar”, “effektives” Verfahren werden wir später vornehmen. Ein weiterer Vorteil einer rekursiven Definition liegt darin, daß man eine Aussage über den neuen Begriff beweisen kann, indem man sie

(1) zunächst für den Anfangsfall und dann

(2) für die aufbauenden Fälle nachweist,

also vom einfachen Fall zum komplizierten übergeht - man spricht dann von einem Beweisverfahren durch *Induktion*. Dieses Prinzip ist bekannt aus der Theorie der natürlichen Zahlen:

**Induktionsprinzip für die natürlichen Zahlen**

*Es sei $E$ eine Eigenschaft mit*

| | |
|---|---|
| (i) $E(0)$, | *Induktionsanfang* |
| (ii) *falls* $E(n)$*, so auch* $E(n')$. | *Induktionsschluß* |

*Dann gilt* $E(n)$ *für alle natürlichen Zahlen n.*

Dabei besteht der Induktionsschluß aus der *Induktionsvoraussetzung* $E(n)$ und der *Induktionsbehauptung* $E(n')$.

Wenn man die gewöhnliche Ordnung auf den natürlichen Zahlen benutzt, kann man das Induktionsprinzip auch in der folgenden Form anwenden:

**(Starkes) Induktionsprinzip für die natürlichen Zahlen**

*Es sei $E$ eine Eigenschaft mit*

(ii') *falls E(m) für alle* $m < n$*, so auch E(n).* *Induktionsschluß*

*Dann gilt* $E(n)$ *für alle natürlichen Zahlen n.*

Diese beiden Induktionsprinzipien sind untereinander und mit dem Minimumsprinzip äquivalent:

**Minimumsprinzip für die natürlichen Zahlen**

*Falls es eine natürliche Zahl m mit der Eigenschaft E(m) gibt,*
*so gibt es eine kleinste derartige Zahl.*

Aussagen über alle Formeln kann man durch Induktion beweisen, indem man ein *Komplexitätsmaß* einführt, d. h. eine Funktion $k$, die allen Formeln eine natürliche Zahl zuordnet, so daß

$$k(\varphi) < k(\neg\varphi), \quad k(\varphi), k(\psi) < k(\varphi * \psi),$$

z. B.: *lz*($\varphi$) = *Anzahl der logischen Verknüpfungen* $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ *in* $\varphi$,
*l*($\varphi$) = *Anzahl aller Zeichen in* $\varphi$ *(Länge von* $\varphi$*).*

Besonders nützlich ist dafür die *Rangfunktion*, rekursiv definiert durch

$$\begin{aligned} \rho(A_i) &= 0, \\ \rho(\neg\varphi) &= \rho(\varphi) + 1, \\ \rho(\varphi * \psi) &= max(\rho(\varphi), \rho(\psi)) + 1, \textit{ wobei } * = \wedge, \vee, \rightarrow, \leftrightarrow . \end{aligned}$$

### 1.1.3 Beweis durch Induktion über den Formelaufbau

*Es sei E eine Eigenschaft mit*

(i) *E(A) gilt für alle Ausagenvariablen A,*

(ii) *gilt* $E(\varphi)$, *so auch* $E(\neg\varphi)$, *und*

(iii) *gelten* $E(\varphi)$ *und* $E(\psi)$, *so auch* $E(\varphi * \psi)$ *für* $* = \wedge, \vee, \rightarrow$ *und* $\leftrightarrow$ .

*Dann gilt* $E(\varphi)$ *für alle a. F.* $\varphi$.

Legt man ein Komplexitätsmaß $k$ zugrunde, so kann man die drei Bedingungen (i) - (iii) ersetzen durch eine einzige:

(ii') *Gilt* $E(\psi)$ *für alle* $\psi$ *mit* $k(\psi) < k(\varphi)$, *so auch* $E(\varphi)$.

Die Rangfunktion haben wir oben durch Rekursion definiert; damit dies zulässig ist, benötigen wir die

### Eindeutige Lesbarkeit von Formeln

*Es sei $\varphi$ eine a. F. Dann ist $\varphi$ entweder eine Aussagenvariable oder läßt sich auf genau eine Weise in der Form*

$$\neg\psi, (\psi \wedge \chi),\ (\psi \vee \chi),\ (\psi \rightarrow \chi) \textit{ oder } (\psi \leftrightarrow \chi)$$

*schreiben.*

Zum *Beweis* benutze man das folgende (technische) Lemma:

### Lemma

*Es sei $\varphi$ eine a. F., w eine endliche Zeichenreihe. Dann gilt:*
*Falls die Zeichenreihe $\varphi w$ eine a. F. ist, so ist w leer, also $\varphi w = \varphi$.*

Dieses Lemma - zugleich ein gutes Beispiel für eine syntaktische Aussage! - beweist man am besten durch Induktion über die Anzahl l($\varphi$w) der Zeichen in $\varphi$w.

Neben der rekursiven Definition eines *Prädikats* (wie im Falle des Formelbegriffs) haben wir im Falle der Rangfunktion eine rekursive Definition einer *Funktion*. Auch derartige Funktionen haben ihr Vorbild in der Zahlentheorie, wo man etwa die Funktion $n!$ rekursiv definiert durch

$$0! = 1, (n+1)! = n!(n+1).$$

Diese Definition liefert nur deshalb eine Funktion, weil für eine Zahl $n \neq 0$ der Vorgänger $m$ von $n$ (mit $m' = n$) eindeutig bestimmt ist. Die entsprechende Voraussetzung für den Formelbegriff haben wir gerade mit der eindeutigen Lesbarkeit der Formeln bewiesen.

### 1.1.4 Definition durch Rekursion über den Formelaufbau

*Zu gegebenen Funktionen $g_0, \ldots, g_5$ (geeigneter Stellenzahl) läßt sich eine Funktion g auf der Menge der a. F. definieren durch:*

$$\begin{aligned}
g(A) &= g_0(A)\\
g(\neg\varphi) &= g_1(g(\varphi))\\
g(\varphi \wedge \psi) &= g_2(g(\varphi), g(\psi))\\
g(\varphi \vee \psi) &= g_3(g(\varphi), g(\psi))\\
g(\varphi \rightarrow \psi) &= g_4(g(\varphi), g(\psi))\\
g(\varphi \leftrightarrow \psi) &= g_5(g(\varphi), g(\psi))
\end{aligned}$$

## 1.2 Semantik der Aussagenlogik

Die *Bedeutung* der aussagenlogischen Verknüpfungen wird üblicherweise durch die Wahrheitstafeln festgelegt - die aussagenlogischen Verknüpfungen werden dadurch zu Funktionen auf den Wahrheitswerten

- W (“wahr”), manchmal auch mit 1 oder V bezeichnet, und
- F (“falsch”), manchmal auch mit 0 oder Λ bezeichnet.

(Wir behandeln hier die 2-wertige Logik, nehmen also an, daß es genau zwei Wahrheitswerte gibt, nämlich W und F.)

### 1.2.1 Wahrheitsfunktionen

1-stellige Wahrheitsfunktionen

| | W | id | $\neg$ | F |
|---|---|---|---|---|
| W | W | W | F | F |
| F | W | F | W | F |

Unter den vier 1-stelligen gibt es nur eine nicht-triviale Wahrheitsfunktion, die die Wahrheitswerte vertauscht und der Negation entspricht. Es gibt $2^{2^2} = 16$ 2-stellige Wahrheitsfunktionen; lassen wir die trivialen weg, so erhalten wir:

2-stellige (nichttriviale) Wahrheitsfunktionen

| | | $\wedge$ | $\vee$ | $\rightarrow$ | $\leftarrow$ | $\nabla$ | $\leftrightarrow$ | $\dashv$ | $\vdash$ | $\downarrow$ | $\uparrow$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| W | W | W | W | W | W | F | W | F | F | F | F |
| W | F | F | W | F | W | W | F | W | F | F | W |
| F | W | F | W | W | F | W | F | F | W | F | W |
| F | F | F | F | W | W | F | W | F | F | W | W |

Neben den bereits eingeführten tritt hier noch das entweder-oder $\nabla$ auf, das weder-noch $\downarrow$ (“nor”, PEIRCE) sowie $\uparrow$ (“nand”, SHEFFER´s stroke).

Weiterhin gibt es bereits $2^{2^3} = 256$ 3-stellige Wahrheitsfunktionen, allgemein $2^{2^n}$ $n$-stellige Wahrheitsfunktionen, auf die wir später noch zurückkommen werden.

Wir wollen nun jeder aussagenlogischen Formel $\varphi(A_1,\dots,A_n)$ in den Aussagenvariablen $A_1,\dots,A_n$ den Wahrheitswert zuordnen, der sich durch Auswertung gemäß den Wahrheitstafeln ergibt. Dieser Wert hängt ab

(1) von den Wahrheitswerten, die die Aussagenvariablen erhalten, sowie davon,

(2) wie $\varphi$ aus den Aussagenvariablen mit Hilfe der aussagenlogischen Verknüpfungen zusammengesetzt ist:

### 1.2.2 Interpretation von aussagenlogischen Formeln

$\varphi$ sei eine aussagenlogische Formel, $V$ sei eine beliebige (u. U. unendliche) Menge von Aussagenvariablen. Man bezeichnet mit

| | |
|---|---|
| $f\colon V \to \{W,F\}$ | *Belegung* von $V$ (mit Wahrheitswerten), |
| $B(V) := \{f \mid f : V \longrightarrow \{W,F\}\}$ | *Menge der Belegungen von V.* |
| $V(\varphi)$ | *Menge der Aussagenvariablen in* $\varphi$ |
| $F(V) := \{\varphi \mid V(\varphi) \subseteq V\}$ | *Menge der Formeln mit Variablen in V.* |

Es sei nun $\varphi$ eine aussagenlogische Formel mit Variablen in $V, f : V \longrightarrow \{W,F\}$ eine Belegung mit $V(\varphi) \subseteq V$. Dann definieren wir

*f(φ) als* **Wahrheitswert** (oder: **Interpretation**) von $\varphi$ *unter der Belegung f*

durch Rekursion über den Aufbau von $\varphi$ wie folgt:

$$\begin{aligned} f(\varphi) &= f(A), \quad falls\ \varphi = A\ eine\ Aussagenvariable\ ist, \\ f(\neg\varphi) &= \neg f(\varphi), \\ f(\varphi * \psi) &= f(\varphi) * f(\psi). \end{aligned}$$

Dabei steht * für die 2-stelligen Operationen. Zu beachten ist, daß auf der linken Seite innerhalb der Klammern die *Symbole* für die aussagenlogischen Verknüpfungen, auf der rechten Seite aber die zugeordneten *Operationen* stehen, die man aus den Wahrheitstafeln entnimmt (und die wir dort der Einfachheit halber auch mit den entsprechenden Symbolen bezeichnet haben).

**Beispiel**

Es sei $\varphi = \neg(A \leftrightarrow C) \wedge \neg D$, also $V(\varphi) = \{A,C,D\}$. Ferner sei $V = \{A,B,C,D\}$, und $f : V \longrightarrow \{W,F\}$ sei definiert durch

$$f(A) = W,\ f(B) = W,\ f(C) = F,\ f(D) = W.$$

Dann erhält man durch sukzessive Berechnung (entsprechend der rekursiven Definition):

| | $A$ | $B$ | $C$ | $D$ | $A \leftrightarrow C$ | $\neg(A \leftrightarrow C)$ | $D$ | $\neg D$ | $\varphi$ |
|---|---|---|---|---|---|---|---|---|---|
| f | W | W | F | W | F | W | W | F | F |

$f(\varphi)$ hängt hier natürlich nicht von $f(B)$ ab, da $B$ nicht in $\varphi$ vorkommt. Allgemein gilt:

**Satz**

*$f(\varphi)$ hängt nur von $f \restriction V(\varphi)$ (und natürlich von $\varphi$) ab, d. h. sind f und g zwei Belegungen, die (mindestens) auf $V(\varphi)$ definiert sind und dort übereinstimmen, so ist $f(\varphi) = g(\varphi)$.*

*Beweis:* durch Induktion über den Formelaufbau von $\varphi$. □

### 1.2.3 Definition der wichtigsten semantischen Begriffe

Es seien $\varphi, \psi \in F(V)$.

$$\begin{array}{rlll}
ag[\varphi] & :\Longleftrightarrow \ \textit{für alle } f \in B(V) & \textit{gilt: } f(\varphi) = W & \varphi \text{ ist } \textbf{allgemeingültig} \\
kd[\varphi] & :\Longleftrightarrow \ \textit{für alle } f \in B(V) & \textit{gilt: } f(\varphi) = F & \varphi \text{ ist } \textbf{kontradiktorisch} \\
erfb[\varphi] & :\Longleftrightarrow \ \textit{es gibt ein } f \in B(V) & \textit{mit: } f(\varphi) = W & \varphi \text{ ist } \textbf{erfüllbar} \\
\varphi \ \textit{äq} \ \psi & :\Longleftrightarrow \ \textit{für alle } f \in B(V) & \textit{gilt: } f(\varphi) = f(\psi) & \varphi \text{ ist } \textbf{äquivalent} \text{ zu } \ \psi \\
\varphi \ \textit{impl} \ \psi & :\Longleftrightarrow \ \textit{für alle } f \in B(V) & \textit{gilt: } f(\varphi) = W \Rightarrow f(\psi) = W & \\
 & \ \ \Longleftrightarrow \ ag[\varphi \to \psi] & & \varphi \ \textbf{impliziert} \ \ \psi
\end{array}$$

**Einfache Beispiele**

$ag[A \vee \neg A]$, $kd[A \wedge \neg A]$, $erfb[A]$, $erfb[\neg A]$, aber $A$ (und auch $\neg A$) sind weder allgemeingültig noch kontradiktorisch.

Es gilt offenbar:

$$\begin{array}{rcl}
ag[\varphi] & \Longleftrightarrow & kd[\neg\varphi] \\
erfb[\varphi] & \Longleftrightarrow & \textit{nicht } kd[\varphi] \\
ag[\neg\varphi] & \Longrightarrow & \textit{nicht } \ ag[\varphi] \quad \text{(aber die Umkehrung gilt i. a. nicht)} \\
ag[\varphi \wedge \psi] & \Longleftrightarrow & ag[\varphi] \ \textit{und} \ ag[\psi] \\
erfb[\varphi \vee \psi] & \Longleftrightarrow & erfb[\varphi] \ \textit{oder} \ erfb[\psi] \\
ag[\varphi] \ \textit{oder} \ ag[\psi] & \Longrightarrow & ag[\varphi \vee \psi] \qquad \text{(i. a. ohne Umkehrung!)} \\
erfb[\varphi \wedge \psi] & \Longrightarrow & erfb[\varphi] \ \textit{und} \ erfb[\psi] \quad \text{(i. a. ohne Umkehrung!)}
\end{array}$$

**Bemerkung**

Die hier eingeführten Prädikate sind entscheidbar. Ein Entscheidungsverfahren liefert (unmittelbar nach Definition) die Wahrheitstafelmethode; dieses Verfahren besitzt jedoch exponentielle Komplexität: Hat $\varphi$ $n$ Variable, so hat $B(V(\varphi))$ bereits $2^n$ Elemente! Für jedes einzelne $f \in B(V(\varphi))$ muß man $f(\varphi)$ berechnen, wobei die Komplexität dieser Berechnung abhängig ist von der Anzahl der logischen Symbole in $\varphi$.

Damit läßt sich die Aussagenlogik kalkülmäßig behandeln - man spricht daher auch vom *Aussagenkalkül*. Dagegen wird sich später zeigen, daß die Allgemeingültigkeit einer prädikatenlogischen Aussage sich i. a. nicht mehr entscheiden läßt; trotzdem spricht man (aus Gründen, die später im Rahmen einer Axiomatisierung deutlicher werden) auch vom *Prädikatenkalkül*.

### 1.2.4 Einige wichtige allgemeingültige Formeln

| | |
|---|---|
| $A \vee \neg A$ | tertium non datur |
| $A \to A$ | Selbstimplikation |
| $\neg(A \wedge \neg A)$ | Gesetz vom Widerspruch |
| $\neg(A \leftrightarrow \neg A)$ | Gesetz vom Widerspruch |
| $\neg A \to (A \to B)$ | regula falsi |
| $A \wedge (A \to B) \to B$ | modus ponens (als Gesetz) |
| $\neg B \wedge (A \to B) \to \neg A$ | modus tollens (als Gesetz) |
| $A \to A \vee B$ | (hintere) Abschwächung |
| $A \wedge B \to A$ | (vordere) Abschwächung |
| $(A \to B) \wedge (B \to C) \to (A \to C)$ | Kettenschluß |
| $(A \vee B) \wedge (A \to C) \wedge (B \to C) \to C$ | Gesetz der Fallunterscheidung |
| $((A \to B) \to A) \to A$ | Formel von Peirce |

Einige dieser allgemeingültigen Formeln ergeben gültige **Implikationen**, wie z. B.

$$A \; impl \; A \; ,$$
$$A \; impl \; B \quad \text{und} \quad B \; impl \; C \Longrightarrow A \; impl \; C \; ,$$
$$A \; impl \; B \quad \text{und} \quad B \; impl \; A \Longrightarrow A \; \ddot{a}q \; B \; .$$

### 1.2.5 Einige wichtige Äquivalenzen

| | |
|---|---|
| $\neg\neg A \; \ddot{a}q \; A$ | Gesetz der doppelten Verneinung |
| $A \to B \; \ddot{a}q \; \neg B \to \neg A$ | Kontraposition |
| $\neg(A \to B) \; \ddot{a}q \; (A \wedge \neg B)$ | |
| $\neg(A \vee B) \; \ddot{a}q \; (\neg A \wedge \neg B)$ | de MORGANsches Gesetz |
| $\neg(A \wedge B) \; \ddot{a}q \; (\neg A \vee \neg B)$ | de MORGANsches Gesetz |
| $A \to (B \to C) \; \ddot{a}q \; A \wedge B \to C$ | Importation / Exportation |
| $A \to B \vee C \; \ddot{a}q \; (A \to B) \vee (A \to C)$ | Distrib. der Implikation “von rechts” |
| $A \to B \wedge C \; \ddot{a}q \; (A \to B) \wedge (A \to C)$ | Distrib. der Implikation “von rechts” |
| $A \vee B \to C \; \ddot{a}q \; (A \to C) \wedge (B \to C)$ | |
| $A \wedge B \to C \; \ddot{a}q \; (A \to C) \vee (B \to C)$ | |

Zu beachten ist der Wechsel von $\vee$ zu $\wedge$ (und umgekehrt) in den beiden letzten Formeln!

### 1.2.6 Boolesche Gesetze

| | |
|---|---|
| $A \cap B = B \cap A$ | Kommutativgesetze |
| $A \cup B = B \cup A$ | |
| $A \cap (B \cap C) = (A \cap B) \cap C$ | Assoziativgesetze |
| $A \cup (B \cup C) = (A \cup B) \cup C$ | |
| $A \cap A = A$ | Idempotenzgesetze |
| $A \cup A = A$ | |
| $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ | Distributivgesetze |
| $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ | |
| $A \cap (B \cup A) = A$ | Absorptionsgesetze |
| $A \cup (B \cap A) = A$ | |
| $A \cup -A = V,\ A \cap -A = \emptyset$ | Komplementgesetze |
| $- -A = A$ | |
| $-\emptyset = V,\ -V = \emptyset$ | |
| $-(A \cap B) = -A \cup -B$ | de MORGANsche Gesetze |
| $-(A \cup B) = -A \cap -B$ | |

Dieses sind die Gesetze einer BOOLEschen Algebra, und zwar in mengentheoretischer Form ausgedrückt. Die entsprechenden aussagenlogischen Gesetze erhält man, wenn man folgende Ersetzungen vornimmt:

| mengentheoretisch | aussagenlogisch |
|---|---|
| $=$ | äq |
| $-$ | $\neg$ |
| $\cap$ | $\wedge$ |
| $\cup$ | $\vee$ |
| $\emptyset$ | $A \wedge \neg A$ |
| $V$ | $A \vee \neg A$ |

Die obigen Gesetze haben wir für Aussagenvariable aufgeschrieben; sie gelten aber allgemeiner für a. F. anstelle der Aussagenvariablen. Diese Tatsache ergibt sich aus der

### 1.2.7 Einsetzungsregel

$$\frac{\varphi(A)}{\varphi[\psi/A]}$$

Damit ist gemeint: Wenn der obere Teil allgemeingültig ist, so auch der untere. Dabei schreiben wir auch $\varphi(A)$, um anzudeuten, daß die Aussagenvariable $A$ in $\varphi(A)$ vorkommt, und mit

$$\varphi[\psi/A]$$

bezeichnen wir die Formel, die aus $\varphi$ wie folgt hervorgeht: Setze in $\varphi$ an *allen* Stellen, an denen die Aussagenvariable $A$ vorkommt, die Formel $\psi$ ein.

**Beispiel**

Es ist $ag[A \vee \neg A]$ und damit auch $ag[\varphi \vee \neg\varphi]$ für jede Formel $\varphi$.

### 1.2.8 Ersetzungsregel

$$\frac{\varphi \leftrightarrow \psi}{\chi \leftrightarrow \chi(\psi/\varphi)}$$

Dabei bezeichnet

$$\chi(\psi/\varphi)$$

die Formel, die aus der Formel $\chi$ wie folgt hervorgeht: Ersetze in $\chi$ an *einigen* (möglicherweise allen) Stellen die Formel $\varphi$ durch die Formel $\psi$.

**Beispiel**

Es ist ag$[\neg\neg A \leftrightarrow A]$, nach der Einsetzungsregel also auch ag$[\neg\neg\varphi \leftrightarrow \varphi]$, und somit erhält man aus einer Formel $\chi$ eine äquivalente Formel, indem man an einigen (oder allen Stellen) $\neg\neg\varphi$ durch $\varphi$ ersetzt.

*Beweis* der Einsetzungsregel:

Es sei $\varphi^* := \varphi[\psi/A]$ mit $V(\varphi) = \{A, A_1, \ldots, A_n\}$ und $V(\psi) = \{B_1, \ldots, B_m\}$, somit: $V(\varphi^*) = \{A_1, \ldots, A_n, B_1, \ldots, B_m\}$.

Für eine beliebige Belegung $g : V(\varphi^*) \longrightarrow \{W, F\}$ ist zu zeigen: $g(\varphi^*) = W$. Dazu definiere man eine Belegung $f : V(\varphi) \longrightarrow \{W, F\}$ durch

$$f(A) = g(\psi),\ f(A_i) = g(A_i)$$

und zeige: $g(\varphi^*) = f(\varphi)$, also $= W$, da ag$[\varphi]$. □

*Beweis* der Ersetzungsregel:

Es sei $\chi^* := \chi(\psi/\varphi)$ und $f : V(\chi) \cup V(\chi^*) \longrightarrow \{W, F\}$ eine beliebige Belegung. Zeige: $f(\chi) = f(\chi^*)$ durch Induktion über den Formelaufbau von $\chi$. □

## 1.3 Normalformen

In diesem Abschnitt werden wir uns mit aussagenlogischen Formeln beschäftigen, die nur die Operationen $\neg, \vee, \wedge$ enthalten und BOOLE**sche Formeln** genannt werden.

### Boolesche Umformung

*Zu jeder aussagenlogischen Formel $\varphi$ gibt es eine äquivalente* BOOLE*sche Formel $\varphi^b$, die dieselben Aussagenvariablen wie $\varphi$ enthält:*

$$\varphi \;\text{ äq }\; \varphi^b, \textit{ wobei } V(\varphi) = V(\varphi^b).$$

*Beweis:* Eliminiere die Symbole $\rightarrow$ und $\leftrightarrow$ durch Umformungen zu äquivalenten Formeln; d. h. ersetze in $\varphi$ jede Teilformel der Form
$\psi \rightarrow \chi$ durch die äquivalente Formel $\neg\psi \vee \chi$,
$\psi \leftrightarrow \chi$ durch die äquivalente Formel $(\neg\psi \vee \chi) \wedge (\neg\chi \vee \psi)$. □

### Bemerkung

Indem man die entsprechenden Äquivalenzen benutzt, kann man auch noch $\vee$ durch $\neg$ und $\wedge$ (bzw. $\wedge$ durch $\neg$ und $\vee$) ausdrücken und erreichen, daß $\varphi^b$ nur noch $\neg$ und $\wedge$ (bzw. $\neg$ und $\vee$) enthält, was im Hinblick auf manche Anwendungen (z. B. Dualitätssatz) aber nicht immer erwünscht ist. Tatsächlich kann man eine aussagenlogische Formel in eine äquivalente umformen, die nur noch eine aussagenlogische Operation (nämlich $\downarrow$ bzw. $\uparrow$) enthält.

### Endliche Konjunktionen und Disjunktionen

$$\begin{array}{ll} \varphi_1 \wedge \varphi_2 \wedge \varphi_3 := \varphi_1 \wedge (\varphi_2 \wedge \varphi_3), & \text{allgemein:} \\ \bigwedge_{i=1,\dots,n} \varphi_i = \varphi_1 & \text{für } n = 1, \\ \bigwedge_{i=1,\dots,n} \varphi_i = \varphi_1 \wedge \bigwedge_{i=2,\dots,n} \varphi_i & \text{für } n > 1, \end{array}$$

und analog definiert man die endlichen Disjunktionen $\bigvee_{i=1,\dots,n} \varphi_i$.

Eine Formel ist in **konjunktiver Normalform (KNF)** gdw sie eine Konjunktion von Disjunktionen ist, wobei die einzelnen Disjunktionsglieder kein $\wedge$ und kein $\vee$ mehr enthalten, d. h. wenn sie von der folgenden speziellen Form ist

$$\bigwedge_{i=1,\dots,n} \bigvee_{j=1,\dots,m_i} \varphi_{ij},$$

wobei jedes $\varphi_{ij}$ eine Aussagenvariable oder eine negierte Aussagenvariable ist, in **disjunktiver Normalform (DNF)** gdw sie von der Form

$$\bigvee_{i=1,\ldots,n} \bigwedge_{j=1,\ldots,m_i} \varphi_{ij}$$

ist, wobei wiederum jedes $\varphi_{ij}$ eine Aussagenvariable oder eine negierte Aussagenvariable ist. (Insbesondere sind $A \vee \neg B$ sowie $A \wedge \neg B$ in konjunktiver und zugleich in disjunktiver Normalform.)

**Satz (Boolesche Normalform)**

*Zu jeder aussagenlogischen Formel $\varphi$ gibt es äquivalente Formeln $\varphi^k$ bzw. $\varphi^d$ in konjunktiver bzw. disjunktiver Normalform, die dieselben Aussagenvariablen wie $\varphi$ enthalten:*

$$\varphi \text{ äq } \varphi^k \text{ äq } \varphi^d, \textit{ wobei}$$
$$V(\varphi) = V(\varphi^k) = V(\varphi^d),\ \varphi^k \textit{ in } \mathrm{KNF}, \varphi^d \textit{ in } \mathrm{DNF}.$$

*Beweis:* Eine disjunktive NF läßt sich aus $\varphi$ mittels folgender Umformungen konstruieren:

(1) $\varphi$ äq $\varphi^b$ für eine BOOLEsche Formel $\varphi^b$.

(2) $\varphi^b$ äq $\varphi^n$ für eine Formel $\varphi^n$, wobei in $\varphi^n$ ein Negationszeichen höchstens vor Aussagenvariablen vorkommt; dazu benutze man folgende Äquivalenzen:

$\neg\neg\psi$ äq $\psi$, $\neg(\psi \vee \chi)$ äq $\neg\psi \wedge \neg\chi$, $\neg(\psi \wedge \chi)$ äq $\neg\psi \vee \neg\chi$ (De Morgan).

(3) eliminiere in $\varphi^n$ $\wedge$ "vor" $\vee$ mittels der Distributivgesetze:

$$\psi \wedge (\chi \vee \delta) \text{ äq } (\psi \wedge \chi) \vee (\psi \wedge \delta)$$
$$(\psi \vee \chi) \wedge \delta \text{ äq } (\psi \wedge \delta) \vee (\chi \wedge \delta)$$

Als Ergebnis erhält man eine disjunktive NF $\varphi^d$, welche zu $\varphi$ äquivalent ist. Die im 2. Schritt erhaltene Formel $\varphi^n$, in welcher das Negationszeichen höchstens vor Aussagenvariablen vorkommt, nennt man auch eine **Negationsnormalform** von $\varphi$. Um eine konjunktive NF zu erhalten, bilde man zuerst wieder eine Negationsnormalform und wende im 3. Schritt die dualen Distributivgesetze an (vertausche $\wedge$ mit $\vee$). (Der Beweis liefert ein effektives Verfahren, um $\varphi$ zu einer äquivalenten konjunktiven bzw. disjunktiven Normalform umzuformen.) □

### 1.3.1 Entscheidungsverfahren für Boolesche Normalformen

(i) $\varphi$ sei in KNF. Dann gilt:
ag[$\varphi$ ] $\Leftrightarrow$ *in jedem Konjunktionsglied von $\varphi$ kommt mindestens eine Aussagenvariable zusammen mit ihrer Negation vor.*

(ii) $\varphi$ sei in DNF. Dann gilt:
kd[$\varphi$ ] $\Leftrightarrow$ *in jedem Disjunktionsglied von $\varphi$ kommt mindestens eine Aussagenvariable zusammen mit ihrer Negation vor.*

**Beispiel**

$$(A \vee D \vee \neg A) \wedge (\neg B \vee A \vee A \vee B) \wedge (\neg A \vee B \vee \neg B \vee A \vee C)$$

ist allgemeingültig, nicht aber $(\neg A \vee B \vee A) \wedge (\neg B \vee A \vee A \vee D)$.

*Beweis* von (i): Es ist

$$\text{ag}[\textstyle\bigwedge_{i=1,\dots,n} \bigvee_j \varphi_{ij}] \iff \textit{für alle } i = 1,\dots,n : \text{ag}[\bigvee_j \varphi_{ij}].$$

Gibt es zu jedem $i$ ein $j$ und ein $k$, so daß $\varphi_{ij} = A$ und $\varphi_{ik} = \neg A$ für eine Aussagenvariable $A$ ist, so ist für jedes $i$ $\bigvee_j \varphi_{ij}$ allgemeingültig, also gilt “$\Leftarrow$”.

Zum Beweis der Umkehrung sei $\varphi = \bigwedge_{i=1,\dots,n} \bigvee_j \varphi_{ij}$ in KNF, aber in (mindestens) *einem* Konjunktionsglied, etwa $\varphi_i = \bigvee_j \varphi_{ij}$, komme *keine* Aussagenvariable mit ihrer Negation vor. Dann können wir eine Belegung $f : V(\varphi) \longrightarrow \{W, F\}$ definieren mit

$$f(A) = \begin{cases} W, & \text{falls } A \text{ in } \varphi_i \text{ negiert vorkommt,} \\ F, & \text{falls } A \text{ in } \varphi_i \text{ unnegiert vorkommt,} \\ \text{beliebig, etwa } W, & \text{falls } A \text{ in } \varphi_i \text{ nicht vorkommt.} \end{cases}$$

Dann wird $f(\varphi_i) = F$ und damit auch $f(\varphi) = F$, also ist $\varphi$ nicht allgemeingültig.
□

Hiermit erhält man ein weiteres Verfahren, die Allgemeingültigkeit einer aussagenlogischen Formel $\varphi$ zu testen: Forme $\varphi$ in eine KNF um und prüfe nach, ob die Bedingung von 1.3.1 (i) erfüllt ist. Auch dieses Entscheidungsverfahren ist von exponentieller Komplexität wegen der Notwendigkeit, bei der Umformung in eine KNF das Distributivgesetz anzuwenden. Dabei werden aus Konjunktionen von Disjunktionen wesentlich längere Disjunktionen von Konjunktionen (und umgekehrt): Das **allgemeine Distributivgesetz** lautet:

$$\bigwedge_{i=1,\ldots,n} \bigvee_{j=1,\ldots,m_i} \varphi_{ij} \text{ äq } \bigvee_{f \in P} \bigwedge_{i=1,\ldots,n} \varphi_{if(i)},$$

wobei $P = \{f \mid f : \{1,\ldots,n\} \longrightarrow \mathbb{N} \textit{ mit } 1 \leq f(i) \leq m_i \textit{ für alle } i = 1,\ldots,n\}$. Falls alle $m_i = 2$ sind, so hat $P$ die Anzahl $2^n$.

Der folgende Satz besagt, daß sich jede $n$-stellige Wahrheitsfunktion mittels der BOOLEschen Operationen $\neg, \vee, \wedge$ darstellen läßt, und aus dem Beweis ergibt sich zugleich eine Darstellung in disjunktiver bzw. konjunktiver Normalform:

### 1.3.2 Boolescher Repräsentationssatz

*Es sei $G$ eine beliebige $n$-stellige Wahrheitsfunktion, also $G : B(V) \longrightarrow \{W,F\}$, wobei $V$ die Menge der Aussagenvariablen $\{A_1,\ldots,A_n\}$ ist.*

*Dann existiert eine* BOOLE*sche Formel $\varphi$ in diesen Ausagenvariablen, d. h. mit $V(\varphi) = V$, deren Auswertung nach den Wahrheitstafeln gerade die Funktion $G$ ergibt:*

$$G(g) = g(\varphi) \textit{ für alle } g \in B(V).$$

*Beweis:* Wir definieren zunächst für jedes $f \in B(V)$ eine Formel $\varphi_f$, die genau bei der Belegung $f$ wahr wird (was wir unten mit (*) zeigen werden):

$$\varphi_f = A_1^{f(A_1)} \wedge \ldots \wedge A_n^{f(A_n)}, \text{ wobei } A_i^W = A_i,\ A_i^F = \neg A_i\ .$$

Es sei $M := \{f \in B(V) \mid G(f) = W\}$. Dann ist

$$\varphi := \bigvee_{f \in M} \varphi_f$$

eine BOOLEsche Formel (in disjunktiver NF), die die Behauptung erfüllt, weil sie gerade für alle $f \in M$ wahr und sonst falsch wird:

$$\begin{array}{llll} & g(\varphi) = W & \Longleftrightarrow g(\varphi_f) = W & \text{für ein } f \in M \\ (*) & & \Longleftrightarrow f = g & \text{für ein } f \in M \\ & & \Longleftrightarrow g \in M & \\ & & \Longleftrightarrow G(g) = W. & \end{array}$$

Zum Beweis von (*) beachte, daß für alle $i = 1,\ldots,n$:

$$\begin{aligned} g(\varphi_f) = W &\Longrightarrow g(A_i^{f(A_i)}) = W \\ &\Longrightarrow [f(A_i) = W \Longrightarrow g(A_i) = W] \text{ und} \\ &\Longrightarrow [f(A_i) = F \Longrightarrow g(A_i) = F], \text{ also } f = g, \end{aligned}$$

und offenbar gilt auch umgekehrt $f(\varphi_f) = W$.

Was passiert, wenn die Funktion $G$ überall den Wert $F$ annimmt? Dann ist die Menge $M$ leer und die zugehörige Formel $\varphi$ eine Disjunktion über eine leere Indexmenge, und eine solche Disjunktion erklärt man als “falsch”. Da wir hierfür keine Konstante in der aussagenlogischen Sprache haben, benutzen wir dafür (wie auch im Satz von BOOLE) die Formel $A_0 \wedge \neg A_0$, ähnlich erklärt man die Konjunktion über eine leere Indexmenge für stets “wahr” (und kann sie mit $A_0 \vee \neg A_0$ gleichsetzen).

Ein alternativer Beweis (welcher zugleich eine konjunktive NF liefert) geht aus von der Menge $N := \{f \in B(V) \mid G(f) = F\}$. □

**Folgerungen**

aus dem obigen Satz (bzw. aus dem Beweis) ergeben sich für

- die Schaltalgebra ,
- den Übergang von einer DNF zu einer äquivalenten KNF,
- die Existenz nur endlich-vieler Formeln $\varphi$ mit Variablen $A_1, \ldots, A_n$, die nicht miteinander äquivalent sind (dagegen gibt es stets unendlich viele zueinander äquivalente Formeln).

### 1.3.3 Das Dualitätsprinzip

Im folgenden betrachten wir nur BOOLEsche Formeln! Wir definieren drei syntaktische Operationen für BOOLEsche Formeln:

$D(\varphi)$ : vertausche $\vee$ mit $\wedge$ in $\varphi$ : **duale Formel** zu $\varphi$,

$N(\varphi)$ : setze vor jede Aussagenvariable in $\varphi$ ein $\neg$-Zeichen,

$N^*(\varphi)$ : setze vor jede Aussagenvariable in $\varphi$ ein $\neg$-Zeichen, falls dort keines steht, sonst streiche es!

**Beispiele**

$$\begin{aligned}
D(\neg(\neg A \vee B) \wedge \neg C) &= \neg(\neg A \wedge B) \vee \neg C \\
N(\neg(\neg A \vee B) \wedge \neg C) &= \neg(\neg\neg A \vee \neg B) \wedge \neg\neg C \\
N^*(\neg(\neg A \vee B) \wedge \neg C) &= \neg(A \vee \neg B) \wedge C \\
DN^*(\neg(\neg A \vee B) \wedge \neg C) &= \neg(A \wedge \neg B) \vee C
\end{aligned}$$

$$\begin{array}{rcl} DN^*(A) & = & \neg A \\ DN^*(A \wedge B) & = & \neg A \vee \neg B \; äq \; \neg(A \wedge B) \\ DN^*(A \vee B) & = & \neg A \wedge \neg B \; äq \; \neg(A \vee B) \end{array}$$

Offensichtlich sind $N(\varphi)$ und $N^*(\varphi)$ miteinander äquivalent, und die letzten drei Formeln kann man verallgemeinern zum Teil a) von

**Satz**

a) $\neg\varphi \; äq \; D(N^*(\varphi))$ **Bildung der Negation**

b) $\varphi \; äq \; \psi \iff D\varphi \; äq \; D\psi$ **Dualitätssatz**

und b) leicht aus a) folgern, was wir hier nicht ausführen. Natürlich kann man eine Aussage negieren, indem man einfach das Negationszeichen davorschreibt; Teil a) des obigen Satzes gibt eine sinnvollere Darstellung, in welcher das Negationszeichen "nach innen" verschoben wird (obwohl man in praktischen Fällen es nicht gleich bis vor die atomaren Teilformeln verschieben wird). Der Dualitätssatz erklärt u. a., weshalb die BOOLEschen Gesetze stets in jeweils doppelter Form auftreten.

## 1.4 Der semantische Folgerungsbegriff: Erfüllbarkeit

Eine Formel ist allgemeingültig, wenn sie uneingeschränkt den Wert W erhält. Diesen Begriff werden wir jetzt relativieren zum Folgerungsbegriff:

### Aussagenlogische Folgerungen

$\varphi$ sei eine Formel, $\mathsf{T}$ eine (möglicherweise unendliche) Menge von Formeln, $V$ eine Menge von Aussagenvariablen mit $\mathsf{T} \subseteq F(V)$, $\varphi \in F(V)$, und schließlich sei $f : V \longrightarrow \{W, F\}$ eine Belegung der Variablen in $V$.

| | | | |
|---|---|---|---|
| $f \models \varphi$ | $:\iff$ | $f(\varphi) = W$ | $f$ **erfüllt** $\varphi$ |
| $f \models \mathsf{T}$ | $:\iff$ | *für alle* $\psi \in \mathsf{T} : f \models \psi$ | $f$ **erfüllt** $\mathsf{T}$ |
| *erfb*$[\varphi]$ | $:\iff$ | *es gibt ein f mit* $f \models \varphi$ | $\varphi$ ist **erfüllbar** |
| *erfb*$[\mathsf{T}]$ | $:\iff$ | *es gibt ein f mit* $f \models \mathsf{T}$ | $\mathsf{T}$ ist **erfüllbar** |
| $\mathsf{T} \models \varphi$ | $:\iff$ | *für alle f : falls* $f \models \mathsf{T}$, *so* $f \models \varphi$ | $\varphi$ **folgt aus** $\mathsf{T}$ |

$\mathsf{T} \models \varphi$ besagt also: alle Belegungen $f$, die $\mathsf{T}$ erfüllen, erfüllen auch $\varphi$,
oder vereinfacht:
$\varphi$ ist immer dann wahr, wenn alle Formeln in $\mathsf{T}$ wahr sind.

Falls $\mathsf{T}$ endlich ist, schreiben wir auch

$$\varphi_1, \dots, \varphi_n \models \varphi \quad \text{für} \quad \{\varphi_1, \dots, \varphi_n\} \models \varphi,$$

und aus $\emptyset \models \varphi$ wird natürlich $\models \varphi$, so daß

$$\models \varphi \iff \mathrm{ag}[\varphi]\,.$$

### Beispiele

(a) $\varphi \models \varphi$, (b) $\varphi \models \varphi \vee \psi$, (c) $\varphi, \psi \models \varphi \wedge \psi$, (d) $\varphi \rightarrow \psi \wedge \delta, \delta \rightarrow \neg\psi \models \neg\varphi$, aber: $A \not\models A \wedge B$.

Im Falle der Ersetzungsregel gilt: $\varphi \leftrightarrow \psi \models \chi \leftrightarrow \chi(\varphi/\psi)$;
im Falle der Einsetzungsregel gilt jedoch i. a. *nicht:* $\varphi(A) \models \varphi[\psi/A]$!

**Satz**

$$\varphi_1, \ldots, \varphi_n \models \varphi \iff \models \varphi_1 \wedge \ldots \wedge \varphi_n \to \varphi \qquad \square$$

Der Folgerungsbegriff für endliche Formelmengen läßt sich somit zurückführen auf den Begriff der Allgemeingültigkeit, der allgemeine Folgerungsbegriff aber auch auf den Erfüllbarkeitsbegriff: Als Verallgemeinerung von

$$\mathrm{ag}[\varphi] \iff \textit{nicht erfb}[\neg\varphi]$$

erhalten wir den

### 1.4.1 Zusammenhang zwischen Folgerung und Erfüllbarkeit

$$\begin{aligned} \mathsf{T} \models \varphi &\iff \textit{nicht erfb}[\mathsf{T} \cup \{\neg\varphi\}] \\ \mathsf{T} \not\models \varphi &\iff \textit{erfb}[\mathsf{T} \cup \{\neg\varphi\}] \end{aligned}$$

**Lemma**

$$\begin{aligned} \textit{erfb}[\mathsf{T}] &\iff \textit{es existiert } \underline{\textit{kein}} \ \varphi \ \textit{mit} \ \mathsf{T} \models \varphi \ \textit{und} \ \mathsf{T} \models \neg\varphi \\ &\iff \textit{es existiert } \underline{\textit{ein}} \ \varphi \ \textit{mit} \ \mathsf{T} \not\models \varphi \end{aligned}$$

**Bemerkungen**

1. Ist $\mathsf{T}$ endlich, so ist entscheidbar, ob $\mathsf{T} \models \varphi$ oder $\mathsf{T} \not\models \varphi$. Es ist aber möglich, daß weder $\mathsf{T} \models \varphi$ noch $\mathsf{T} \models \neg\varphi$ gilt (wenn $\mathsf{T}$ nämlich zu wenig Information über die Variablen in $\varphi$ enthält, z. B. $A \not\models B$ und $A \not\models \neg B$ für zwei verschiedene Variable $A$ und $B$)! Denn es gilt zwar

   $\mathsf{T} \models \neg\varphi \implies \mathsf{T} \not\models \varphi$ (für erfüllbares $\mathsf{T}$), aber nicht umgekehrt!

2. Ist $\mathsf{T}$ endlich, so gibt es eine effektive Aufzählung der Folgerungsmenge $\{\varphi \mid \mathsf{T} \models \varphi\}$.

3. Allgemeiner: Gibt es eine effektive Aufzählung von $\mathsf{T}$, so gibt es auch eine effektive Aufzählung der Folgerungsmenge $\{\varphi \mid \mathsf{T} \models \varphi\}$.

## 1.5 Der syntaktische Folgerungsbegriff: Beweisbarkeit

### 1.5.1 Axiomensysteme und Beweise

Ein (formales) Axiomensystem $\mathcal{A}$ wird bestimmt durch

1. die **Sprache** von $\mathcal{A}$, gegeben durch eine (i. a. abzählbare) Menge von Symbolen, dem **Alphabet** der Sprache von $\mathcal{A}$,

   z. B. $A_0, A_1, \dots, \neg, \vee, \wedge, \rightarrow, \leftrightarrow, (,)$ im Falle der Aussagenlogik,

2. eine Menge von **Formeln** als bestimmte endliche Folgen von Symbolen, z. B. den aussagenlogischen Formeln,

3. eine Menge von **Axiomen**, wobei jedes Axiom eine Formel ist,

   z. B. $\varphi \vee \neg\varphi$, $\varphi \rightarrow \neg\neg\varphi$,

4. eine Menge von **Regeln** der Form $\dfrac{\varphi}{\psi}$ oder $\dfrac{\varphi,\ \psi}{\delta}$,

   z. B. $\dfrac{\varphi}{\neg\neg\varphi}$, modus ponens: $\dfrac{\varphi,\ \varphi \rightarrow \psi}{\psi}$.

Es sei nun $\mathcal{A}$ ein Axiomensystem, $\varphi$ eine Formel, $\mathsf{T}$ eine Menge von Formeln (der Sprache von $\mathcal{A}$). Ein ($\mathcal{A}$-)**Beweis** von $\varphi$ aus $\mathsf{T}$ ist eine endliche Folge $(\varphi_1, \dots, \varphi_n)$ von Formeln, so daß

(B1) $\varphi_n = \varphi$,

(B2) für jedes $m \leq n$ gilt:

(1) $\varphi_m$ ist Axiom von $\mathcal{A}$ *oder*

(2) $\varphi_m$ ist Formel aus $\mathsf{T}$ *oder*

(3) es gibt ein $i < m$, so daß $\varphi_m$ aus $\varphi_i$ hervorgeht durch Anwendung einer Regel von $\mathcal{A}$ *oder*

(4) es gibt $i, j < m$, so daß $\varphi_m$ aus $\varphi_i$ und $\varphi_j$ hervorgeht durch Anwendung einer Regel von $\mathcal{A}$ mit 2 Prämissen.

Wir schreiben (analog zum semantischen Folgerungsbegriff):

| | | |
|---|---|---|
| $\mathsf{T} \vdash_{\mathcal{A}} \varphi :\iff$ | es gibt einen $\mathcal{A}$-Beweis von $\varphi$ aus $\mathsf{T}$ | $\varphi$ ist aus $\mathsf{T}$ **beweisbar**, |
| $\vdash_{\mathcal{A}} \varphi :\iff$ | $\emptyset \vdash \varphi$ | $\varphi$ ist (in $\mathcal{A}$ ) **beweisbar**. |

Statt $\mathsf{T} \vdash_{\mathcal{A}} \varphi$ werden wir auch kurz $\mathsf{T} \vdash \varphi$ schreiben, wenn das Axiomensystem aus dem Zusammenhang bekannt (oder unwesentlich) ist, statt "$\varphi$ ist aus $\mathsf{T}$ beweisbar" auch "$\varphi$ folgt aus $\mathsf{T}$" oder "$\varphi$ ist aus $\mathsf{T}$ ableitbar" sagen.

### 1.5.2 Korrektheit und Vollständigkeit

| | | | |
|---|---|---|---|
| $\mathcal{A}$ **korrekt** | $:\iff$ | *für alle* $\varphi$: | $\vdash_{\mathcal{A}} \varphi \Rightarrow\ \models \varphi$, |
| $\mathcal{A}$ **korrekt bzgl. Folgerungen** | $:\iff$ | *für alle* $\varphi, \mathsf{T}$: | $\mathsf{T} \vdash_{\mathcal{A}} \varphi \Rightarrow \mathsf{T} \models \varphi$, |
| $\mathcal{A}$ (syntaktisch) **widerspruchsfrei** | $:\iff$ | *es gibt ein* $\varphi$ *mit* | $\not\vdash_{\mathcal{A}} \varphi$, |
| $\mathcal{A}$ (semantisch) **vollständig** | $:\iff$ | *für alle* $\varphi$: | $\models \varphi \Rightarrow\ \vdash_{\mathcal{A}} \varphi$. |
| $\mathcal{A}$ **vollständig bzgl. Folgerungen** | $:\iff$ | *für alle* $\varphi, \mathsf{T}$: | $\mathsf{T} \models \varphi \Rightarrow \mathsf{T} \vdash_{\mathcal{A}} \varphi$. |

In einem korrekten Axiomensystem sind alle beweisbaren Aussagen allgemeingültig, für ein vollständiges Axiomensystem gilt die Umkehrung. Vollständigkeit und Korrektheit zusammen besagen also, daß genau die allgemeingültigen Aussagen beweisbar sind.

- Ein *korrektes* Axiomensystem ist *widerspruchsfrei*: die Aussagenvariable A ist nicht allgemeingültig, kann dann also auch nicht beweisbar sein.

- In einem *widerspruchsvollen* (d. h. nicht widerspruchsfreien) Axiomensystem ist *alles* beweisbar;

- ein *widerspruchsfreies* Axiomensystem (mit modus ponens als Regel) enthält *keinen Widerspruch* in dem Sinne, daß keine Formel und zugleich auch ihre Negation beweisbar ist:

**Bemerkung**

Enthält $\mathcal{A}$ den modus ponens als Regel und gilt

$$\vdash_{\mathcal{A}} \varphi \to (\neg\varphi \to \psi) \text{ oder } \vdash_{\mathcal{A}} \neg\varphi \to (\varphi \to \psi), \text{ so:}$$

$$\mathcal{A} \quad \text{widerspruchsfrei} \iff \text{es gibt keine Formel } \varphi \text{ mit } \vdash_{\mathcal{A}} \varphi \text{ und } \vdash_{\mathcal{A}} \neg\varphi.$$

**Satz**

$\mathsf{T}$ *und* $\mathsf{S}$ *seien Formelmengen,* $\varphi$ *eine Formel.*

1. *Ist* $\varphi \in \mathsf{T}$ *oder* $\varphi$ *Axiom von* $\mathcal{A}$, *so gilt* $\mathsf{T} \vdash_{\mathcal{A}} \varphi$.

2. (**Endlichkeitssatz** für $\vdash$)

   *Falls* $\mathsf{T} \vdash_{\mathcal{A}} \varphi$, *so existiert ein endliches* $\mathsf{T}_0 \subseteq \mathsf{T}$ *mit* $\mathsf{T}_0 \vdash_{\mathcal{A}} \varphi$.

3. (Transitivität von $\vdash$) *Falls* $\mathsf{T} \vdash_{\mathcal{A}} \varphi$ *und* $\mathsf{S} \vdash_{\mathcal{A}} \psi$ *für jedes* $\psi \in \mathsf{T}$, *so* $\mathsf{S} \vdash_{\mathcal{A}} \varphi$.

4. $\mathcal{A}$ *enthalte den modus ponens als Regel. Dann gilt:*
   $\mathsf{T} \vdash_{\mathcal{A}} \varphi$ *und* $\mathsf{T} \vdash_{\mathcal{A}} \varphi \to \psi \Longrightarrow \mathsf{T} \vdash_{\mathcal{A}} \psi$. $\square$ .

### 1.5.3 Einige Axiomensysteme der Aussagenlogik

Es gibt zahlreiche Beispielen für Axiomensysteme der Aussagenlogik mit unterschiedlichen Vorzügen und Nachteilen, wir erwähnen nur die folgenden:

#### (a) Hilbert-Bernays

- Sprache: $A, B, C, \dots, \neg, \vee, \wedge, \to, \leftrightarrow, (,)$
- Axiome:

  1. $A \to (B \to A)$
     $(A \to (A \to B)) \to (A \to B)$
     $(A \to B) \to ((B \to C) \to (A \to C))$
  2. $A \wedge B \to A$
     $A \wedge B \to B$
     $(A \to B) \to ((A \to C) \to (A \to B \wedge C))$
  3. $A \to A \vee B$
     $B \to A \vee B$
     $(A \to C) \to ((B \to C) \to (A \vee B \to C))$
  4. $(A \leftrightarrow B) \to (A \to B)$
     $(A \leftrightarrow B) \to (B \to A)$
     $(A \to B) \to ((B \to A) \to (A \leftrightarrow B))$ (bis hier: *Positive Logik*)
  5. $(A \to B) \to (\neg B \to \neg A)$ (bis hier: *Minimalkalkül*)
     $A \to \neg\neg A$ (bis hier: *Intuitionistische Logik*)
     $\neg\neg A \to A$ (*klassische Logik*)

- Regeln:

  Einsetzungsregel: $\dfrac{\varphi(A)}{\varphi[\psi/A]}$ und modus ponens: $\dfrac{\varphi,\ \varphi \to \psi}{\psi}$

Interessant ist hier die Gruppierung der Axiome nach den einzelnen aussagenlogischen Operationen, zudem ist die Anordnung der letzten Axiome so gewählt, daß man durch Weglassen verschiedene Teilsysteme der Aussagenlogik erhält, die sich durch ihre jeweiligen Anforderungen an die Negation unterscheiden.

**(b) Shoenfield**

- Sprache: $A, B, C, \ldots, \neg, \vee, (,)$

  Die übrigen aussagenlogischen Operationen werden wie üblich definiert, endliche Folgen von Disjunktionen und Implikationen werden durch Klammerung von rechts definiert:

$$\begin{aligned} \varphi_1 \to \varphi_2 \ldots \to \varphi_n \quad &: \leftrightarrow \varphi_1 \to (\varphi_2 \to (\ldots \varphi_n)), \\ \varphi_1 \vee \varphi_2 \ldots \vee \varphi_n \quad &: \leftrightarrow \varphi_1 \vee (\varphi_2 \vee (\ldots \varphi_n)). \end{aligned}$$

- Axiome: $\neg\varphi \vee \varphi$ (tertium non datur)
- Regeln:

$$\begin{array}{llll} \text{Expansion:} & \dfrac{\psi}{\varphi \vee \psi} & \text{Assoziativität:} & \dfrac{\varphi \vee (\psi \vee \delta)}{(\varphi \vee \psi) \vee \delta} \\ \text{Kürzung:} & \dfrac{\varphi \vee \varphi}{\varphi} & \text{Schnitt:} & \dfrac{\varphi \vee \psi, \neg\varphi \vee \delta}{\psi \vee \delta} \end{array}$$

Die Axiome sind hier als *Formelschemata* aufgeschrieben, so daß keine Einsetzungsregel erforderlich ist.

**(c) Nicod**

Dieses Axiomensystem kommt allein mit einer aussagenlogischen Verknüpfung aus, nämlich mit $\uparrow$ (wobei $A \uparrow B$ bedeutet: *nicht A oder nicht B*), außerdem wird auch nur nur ein Axiomenschema und nur eine Regel benötigt:

$$(\alpha \uparrow (\beta \uparrow \gamma)) \uparrow \{[\delta \uparrow (\delta \uparrow \delta)] \uparrow [(\varepsilon \uparrow \beta) \uparrow ((\alpha \uparrow \varepsilon) \uparrow (\alpha \uparrow \varepsilon))]\}$$

$$\frac{\alpha, \alpha \uparrow (\beta \uparrow \gamma)}{\gamma}$$

Das ist wohl das allerkürzeste (wenn auch nicht besonders intuitive) Axiomensystem. Hat ein Axiomensystem wenig Grundsymbole und/oder wenig Axiome und Regeln, so läßt sich seine Korrektheit i. a. sehr schnell zeigen, dagegen werden Beweise *in* einem solchen System sehr lang und umständlich, und die Vollständigkeit ist dann entsprechend mühsam zu zeigen.

Alle diese Axiomensysteme sind korrekt, widerspruchsfrei und vollständig, jedoch nicht bezüglich Folgerungen, wenn sie die Einsetzungsregel enthalten. Um diese Regel zu vermeiden, muß man die Axiome (wie in den letzten beiden Axiomensystemen) als Formelschemata schreiben!

Die Korrektheit beweist man in allen Fällen durch *Induktion über die Länge eines Beweises*, d. h. man zeigt

- die Axiome sind allgemeingültig und
- sind die Prämissen einer Regel allgemeingültig, so auch die Konklusion.

Gilt außerdem:

- aus der (den) Prämisse(n) einer Regel folgt die Konklusion,

so sind die Axiomensysteme auch korrekt bzgl. Folgerungen; modus ponens und die übrigen betrachteten Regeln - außer der Einsetzungsregel! - erfüllen diese Bedingung. So kann man leicht nachweisen, daß das Axiomensystem von SHOENFIELD korrekt bezüglich Folgerungen ist.

## 1.6 Vollständigkeit und Kompaktheit

Die Vollständigkeit eines Axiomensystems ist i. a. schwieriger als seine Korrektheit zu beweisen, und wie es auch kein irgendwie “kanonisches” Axiomensystem für die Aussagenlogik gibt, so gibt es auch keine Standardmethode zum Nachweis der Vollständigkeit, vielmehr muß man die Beweismethode dem Charakter des jeweiligen Axiomensystems anpassen. Hier wollen wir als Beispiel das Axiomensystem von SHOENFIELD behandeln. Der Beweisbegriff $\vdash$ bezieht sich also auf dieses System; die Ergebnisse gelten jedoch (mit möglicherweise anderen Beweisen) auch für andere vollständige Axiomensysteme der Aussagenlogik!

**Kommutativität**

$$\varphi \vee \psi \vdash \psi \vee \varphi$$

*Beweis:* Aus der Voraussetzung $\varphi \vee \psi$ und dem Axiom $\neg\varphi \vee \varphi$ erhält man mittels der Schnittregel: $\psi \vee \varphi$. □

Das bedeutet also, daß man

$$\frac{\varphi \vee \psi}{\psi \vee \varphi}$$

als Regel hinzunehmen kann (man sagt auch, daß diese Regel *ableitbar* sei).

**Modus Ponens**

$$\varphi, \varphi \to \psi \vdash \psi$$

d. h. auch der modus ponens

$$\frac{\varphi, \varphi \to \psi}{\psi}$$

ist eine ableitbare Regel.

*Beweis:* Aus

$$\begin{array}{ll} \varphi & \text{(Voraussetzung)} \\ \psi \vee \varphi & \text{(Expansion) erhalten wir mit Komm. und der Vor. } \neg\varphi \vee \psi \\ \dfrac{\varphi \vee \psi,\ \neg\varphi \vee \psi}{\psi \vee \psi} & \text{(Schnitt)} \\ \psi \quad \text{(Kürzung)} & \end{array}$$

□

Die Kommutativregel, Expansionsregel und die Assoziativregel lassen sich verallgemeinern, wobei wir mit einem Spezialfall beginnen:

**Lemma**

*Für $1 \leq i < j \leq n$ gilt:* $\quad \varphi_i \vee \varphi_j \vdash \varphi_1 \vee \ldots \vee \varphi_n$.

*Beweis* durch Induktion über $n$, wobei man $n > 2$ annehmen kann: Setze $\varphi := \varphi_3 \vee \ldots \vee \varphi_n$. Dann ist zu zeigen: $\varphi_i \vee \varphi_j \vdash \varphi_1 \vee (\varphi_2 \vee \varphi)$.

Fall 1: $i \geq 2$. Dann gilt $\varphi_i \vee \varphi_j \vdash \varphi_2 \vee \varphi$ nach Induktionsvoraussetzung (da $\varphi_1$ weggelassen wurde, handelt es sich nur noch um $n-1$ Formeln). Also gilt nach der Expansionsregel: $\varphi_i \vee \varphi_j \vdash \varphi_1 \vee (\varphi_2 \vee \varphi)$.

Fall 2: $i = 1, j \geq 3$. Dann ist aus der Voraussetzung $\varphi_i \vee \varphi_j$ beweisbar:

| | |
|---|---|
| $\varphi_1 \vee \varphi$ | nach Ind.vor. für die $n-1$ Formeln ohne $\varphi_2$ |
| $\varphi \vee \varphi_1$ | nach Komm. |
| $\varphi_2 \vee (\varphi \vee \varphi_1)$ | mit der Expansionsregel |
| $(\varphi_2 \vee \varphi) \vee \varphi_1$ | mit der Assoziativregel |
| $\varphi_1 \vee (\varphi_2 \vee \varphi)$ | nach Komm. |

Fall 3: $i = 1, j = 2$. Dann ist aus der Voraussetzung $\varphi_i \vee \varphi_j$, d. h. $\varphi_1 \vee \varphi_2$ beweisbar:

| | |
|---|---|
| $\varphi \vee (\varphi_1 \vee \varphi_2)$ | mit der Expansionsregel |
| $(\varphi \vee \varphi_1) \vee \varphi_2$ | mit der Assoziativregel |
| $\varphi_2 \vee (\varphi \vee \varphi_1)$ | nach Komm. |
| $(\varphi_2 \vee \varphi) \vee \varphi_1$ | mit der Assoziativregel |
| $\varphi_1 \vee (\varphi_2 \vee \varphi)$ | nach Komm. |

□

Mit ähnlichen kombinatorischen Überlegungen zeigt man:

### 1.6.1 Verallgemeinerte Expansion

*Für* $n, m \geq 1, i_1, \dots i_m \in \{1, \dots, n\}$ *gilt:* $\quad \varphi_{i_1} \vee \dots \vee \varphi_{i_m} \vdash \varphi_1 \vee \dots \vee \varphi_n$.

*Beweis* durch Induktion über $m$:

Fall 1: $m = 1, i_1 = i$. Dann gilt $\varphi_i$ nach Induktionsvoraussetzung, also:

| | |
|---|---|
| $(\varphi_{i+1} \vee \dots \vee \varphi_n) \vee \varphi_i$ | mit der Expansionsregel |
| $\varphi_i \vee \varphi_{i+1} \vee \dots \vee \varphi_n$ | nach Komm. |
| $\varphi_1 \vee \dots \vee \varphi_n$ | mit der Expansionsregel $(i-1)$-mal. |

Fall 2: $m = 2$. Dieser Fall ergibt sich aus dem vorigen Lemma zusammen mit Komm.

Fall 3: $m \geq 3$. Setze $\varphi := \varphi_1 \vee \dots \vee \varphi_n$. Aus der Voraussetzung

| | |
|---|---|
| $\varphi_{i_1} \vee \dots \vee \varphi_{i_m}$ | folgt |
| $(\varphi_{i_1} \vee \varphi_{i_2}) \vee \dots \vee \varphi_{i_m}$ | mit der Assoziativregel |
| $(\varphi_{i_1} \vee \varphi_{i_2}) \vee \varphi$ | nach Ind.vor. für $m-1$ Formeln |
| $(\varphi \vee \varphi_{i_1}) \vee \varphi_{i_2}$ | mit Komm.- und Assoziativregel |
| $(\varphi \vee \varphi_{i_1}) \vee \varphi$ | nach Ind.vor. für 2 Formeln $(m = 2)$ |

| | |
|---|---|
| $(\varphi \vee \varphi) \vee \varphi_{i_1}$ | mit der Komm.- und Assoziativregel |
| $(\varphi \vee \varphi) \vee \varphi$ | nach Induktionsvor. für 2 Formeln ($m = 2$) |
| $\varphi$ | nach Assoziativregel und Fall 1. |

Dabei zählt bei der Induktionsvoraussetzung für $m-1$ Formeln die Formel $(\varphi_{i_1} \vee \varphi_{i_2})$ als eine Formel! □

Mit der Kommutativregel erhalten wir hieraus auch die Umkehrung der Assoziativregel:

$$\frac{(\varphi \vee \psi) \vee \delta}{\varphi \vee (\psi \vee \delta)}$$

### 1.6.2 Lemma über die Negation

(i) $\varphi \vee \psi \vdash \neg\neg\varphi \vee \psi$

(ii) $\neg\neg\varphi \vee \psi \vdash \varphi \vee \psi$

(iii) $\neg\varphi \vee \delta, \neg\psi \vee \delta \vdash \neg(\varphi \vee \psi) \vee \delta$

Den *Beweis* überlassen wir als Übungsaufgabe.
Nunmehr können wir den (schwachen) Vollständigkeitssatz beweisen als

### 1.6.3 Tautologiesatz

(i) $\models \varphi \implies \vdash \varphi$

(ii) $\varphi_1, \ldots, \varphi_n \models \varphi \implies \varphi_1, \ldots, \varphi_n \vdash \varphi$

*Beweis:* Zunächst zeigen wir, daß (ii) bereits aus (i) folgt: Sei $\varphi_1, \ldots, \varphi_n \models \varphi$. Dann gilt $\models \varphi_1 \to \ldots \to \varphi_n \to \varphi$, also nach (i):

| | |
|---|---|
| $\vdash \varphi_1 \to \ldots \to \varphi_n \to \varphi$ | und also auch |
| $\varphi_1 \vdash \varphi_1 \to \ldots \to \varphi_n \to \varphi,$ | woraus mit modus ponens folgt |
| $\varphi_1 \vdash \varphi_2 \to \ldots \to \varphi_n \to \varphi,$ | ebenso: |
| $\varphi_1, \varphi_2 \vdash \varphi_3 \to \ldots \to \varphi_n \to \varphi,$ | usw. bis |
| $\varphi_1, \ldots, \varphi_n \vdash \varphi.$ | |

Der Beweis von (i) erfolgt durch eine raffiniert angesetzte Induktion: Wir zeigen

$$\models \varphi_1 \vee \ldots \vee \varphi_n \implies \vdash \varphi_1 \vee \ldots \vee \varphi_n$$

durch Induktion über $\Sigma_{i=1,\ldots,n}\, lz(\varphi_i)$, wobei $lz(\varphi)$ = Anzahl der logischen Zeichen in $\varphi$ ist:

1. Fall: Jedes $\varphi_i$ ist eine Aussagenvariable oder eine negierte Aussagenvariable. Falls $\models \varphi_1 \vee \ldots \vee \varphi_n$, so muß ein $\varphi_i = \neg\varphi_j$ für ein $j$ sein. Es ist dann aber $\varphi_i \vee \varphi_j$ ein Axiom, also beweisbar, und damit nach 1.6.1 auch $\vdash \varphi_1, \ldots, \varphi_n$.

Falls der 1. Fall nicht vorliegt, so ist mindestens ein $\varphi_i$ keine Aussagenvariable und auch keine negierte Aussagenvariable; durch Vertauschen können wir annehmen, daß diese Voraussetzung auf $\varphi_1$ zutrifft.

2. Fall: $\varphi_1 = \psi \vee \delta$ für ein $\psi$ bzw. $\delta$. Dann folgt aus der Annahme: $\models \psi \vee (\delta \vee \ldots \vee \varphi_n)$, also gilt nach Induktionsvoraussetzung (hier liegt der Trick des Beweises!) auch $\vdash \psi \vee (\delta \vee \ldots \vee \varphi_n)$ und nach der Assoziativregel dann auch $\vdash \varphi_1 \vee \ldots \vee \varphi_n$.

3. Fall: $\varphi_1 = \neg\neg\psi$ für ein $\psi$. Dann folgt aus der Annahme $\models \psi \vee \varphi_2 \vee \ldots \vee \varphi_n$, also nach Induktionsvoraussetzung auch $\vdash \psi \vee \varphi_2 \vee \ldots \vee \varphi_n$, und nach Lemma 1.6.2 (i) gilt dann $\vdash \varphi_1 \vee \ldots \vee \varphi_n$.

4. Fall: $\varphi_1 = \neg(\psi \vee \delta)$ für ein $\psi$ bzw. $\delta$. Dann ist unter der Annahme

| | | | |
|---|---|---|---|
| $\models \neg(\psi \vee \delta) \vee \varphi_2 \vee \ldots \vee \varphi_n$, | | | auch |
| $\models \neg\psi \vee \varphi_2 \vee \ldots \vee \varphi_n$ | und | $\models \neg\delta \vee \varphi_2 \vee \ldots \vee \varphi_n$, | und somit |
| $\vdash \neg\psi \vee \varphi_2 \vee \ldots \vee \varphi_n$ | und | $\vdash \neg\delta \vee \varphi_2 \vee \ldots \vee \varphi_n$ | nach Ind.vor., woraus |
| $\vdash \varphi_1 \vee \ldots \vee \varphi_n$ | mit Lemma 1.6.2 (iii) folgt. | | |

□

Mit Hilfe des Tautologiesatzes kann man nun beweisen:

### 1.6.4 Deduktionstheorem

$$\mathsf{T} \vdash \varphi \rightarrow \psi \iff \mathsf{T} \cup \{\varphi\} \vdash \psi$$

Wir werden dieses rein syntaktische Ergebnis später im Rahmen der Prädikatenlogik auch rein syntaktisch beweisen.

### 1.6.5 Widerspruchsfreiheit und Vollständigkeit von Theorien

*Widerspruchsfreiheit (Konsistenz)* und *Vollständigkeit* haben wir als syntaktische Begriffe in 1.5.2 für Axiomensysteme definiert; der erste Begriff wird entsprechend auf Formelmengen (später: Theorien) übertragen. Aufpassen muß man dagegen bei dem Begriff der *Vollständigkeit*: Dieser wird im Falle von Formelmengen eine andere Bedeutung haben als für Axiomensysteme.

$$\begin{aligned}
\mathsf{T}\ \textbf{konsistent}: &\iff \text{es gibt } \textit{ein}\ \varphi \text{ mit } \mathsf{T} \nvdash \varphi \\
&\iff \text{es gibt } \textit{kein}\ \varphi \text{ mit } \mathsf{T} \vdash \varphi \text{ und } \mathsf{T} \vdash \neg\varphi \\
\mathsf{T}\ \textbf{vollständig}: &\iff \text{für alle } \varphi \text{ gilt: } \mathsf{T} \vdash \varphi \text{ oder } \mathsf{T} \vdash \neg\varphi.
\end{aligned}$$

Somit ist eine Formelmenge

$$\mathsf{T} \text{ vollständig und konsistent } \iff \text{für alle } \varphi \text{ gilt: } \underline{\text{entweder}}\ \mathsf{T} \vdash \varphi\ \underline{\text{oder}}\ \mathsf{T} \vdash \neg\varphi.$$

Gleichbedeutend damit ist die Aussage, daß die *Menge der Folgerungen aus* $\mathsf{T}$

$$C(\mathsf{T}) := \{\sigma \mid \mathsf{T} \vdash \sigma\}$$

**maximal konsistent** ist, d. h. $C(\mathsf{T})$ ist konsistent, besitzt aber keine echte konsistente Erweiterung.

Eine Theorie ist **inkonsistent** (oder **widerspruchsvoll**) gdw. sie nicht konsistent ist; ähnlich wie im Fall der semantischen Analoga (1.4.1) hängen Beweisbarkeit und (In-)konsistenz miteinander zusammen:

### 1.6.6 Lemma über Beweisbarkeit und Konsistenz

$$\begin{aligned}
(i)\quad \mathsf{T} \nvdash \varphi &\iff \mathsf{T} \cup \{\neg\varphi\}\ \textit{konsistent},\\
(ii)\quad \mathsf{T} \vdash \varphi &\iff \mathsf{T} \cup \{\neg\varphi\}\ \textit{inkonsistent}.
\end{aligned}$$

*Beweis:* Gilt $\mathsf{T} \vdash \varphi$, so erst recht $\mathsf{T} \cup \{\neg\varphi\} \vdash \varphi$ und natürlich auch $\mathsf{T} \cup \{\neg\varphi\} \vdash \neg\varphi$, also ist $\mathsf{T} \cup \{\neg\varphi\}$ inkonsistent.

Ist umgekehrt $\mathsf{T} \cup \{\neg\varphi\}$ inkonsistent, so ist alles beweisbar, insbesondere auch $\mathsf{T} \cup \{\neg\varphi\} \vdash \varphi$ und nach dem Deduktionstheorem $\mathsf{T} \vdash \neg\varphi \rightarrow \varphi$, woraus $\mathsf{T} \vdash \varphi$ folgt. Damit haben wir (ii) gezeigt (und damit auch (i)). $\square$

Den verallgemeinerten Vollständigkeitssatz werden wir nun in der Form

$$\boxed{\textbf{jede konsistente Theorie ist erfüllbar}}$$

beweisen. Dazu werden wir in zwei Schritten zeigen:

- jede konsistente Menge $\mathsf{T}$ läßt sich zu einer vollständigen Menge erweitern,
- eine vollständige Menge ist (in kanonischer) Weise erfüllbar.

### 1.6.7 Vervollständigung

*Zu jeder konsistenten Formelmenge* $\mathsf{T}$ *gibt es eine vollständige Erweiterung* $\mathsf{T}_V \supseteq \mathsf{T}$.

*Beweis:* Es sei $\varphi_0, \ldots, \varphi_n, \ldots$ eine Aufzählung aller Formeln. Durch Induktion erweitern wir $\mathsf{T}$, indem wir $\mathsf{T}_0 = \mathsf{T}$ setzen und

$$\mathsf{T}_{n+1} = \begin{cases} \mathsf{T}_n & \text{falls } \mathsf{T}_n \vdash \varphi_n, \\ \mathsf{T}_n \cup \{\neg\varphi_n\} & \text{sonst.} \end{cases}$$

Nach obigem Lemma 1.6.6 (i) bleibt $\mathsf{T}_{n+1}$ auch im "sonst"-Fall konsistent, so daß alle $\mathsf{T}_n$ konsistent sind und damit auch die Vereinigung

$$\mathsf{T}_V := \bigcup_{n \in \mathbb{N}} \mathsf{T}_n$$

(nach dem Endlichkeitssatz für die Beweisbarkeit, s. 1.5.2). Da die Formel $\varphi_n$ oder ihre Negation spätestens im $(n+1)$-ten Schritt zu $\mathsf{T}$ hinzugefügt werden, ist $\mathsf{T}_V$ auch vollständig. □

### 1.6.8 Verallgemeinerter Vollständigkeitssatz

$$\begin{array}{llcll} (i) & \mathsf{T} \models \varphi & \Longrightarrow & \mathsf{T} \vdash \varphi, & also \\ & \mathsf{T} \models \varphi & \Longleftrightarrow & \mathsf{T} \vdash \varphi. & \\ (ii) & \mathsf{T}\ konsistent & \Longrightarrow & \mathsf{T}\ hat\ ein\ Modell, & also \\ & \mathsf{T}\ konsistent & \Longleftrightarrow & \mathsf{T}\ hat\ ein\ Modell. & \end{array}$$

*Beweis* von (i) mittels (ii):

$$\begin{array}{lll} \mathsf{T} \nvdash \varphi & \Longrightarrow \mathsf{T} \cup \{\neg\varphi\}\ konsistent & \text{nach 1.6.6 (i)} \\ & \Longrightarrow \mathsf{T} \cup \{\neg\varphi\}\ erfüllbar & \text{nach (ii)} \\ & \Longrightarrow \mathsf{T} \not\models \varphi. & \end{array}$$

(Umgekehrt kann man aus (i) auch leicht (ii) erhalten!)

*Beweis von* (ii): Es sei $\mathsf{T}$ konsistent, $\mathsf{T}_V$ eine vollständige Erweiterung von $\mathsf{T}$. Wir definieren dann eine Belegung $f$ von allen Aussagenvariablen durch

$$f(A) = W \iff \mathsf{T}_V \vdash A.$$

Dann gilt allgemeiner für alle Formeln $\varphi$:

$$(*) \quad f \models \varphi \iff \mathsf{T}_V \vdash \varphi.$$

Diese Behauptung beweist man leicht durch Induktion über $\rho(\varphi)$:

1. Ist $\varphi$ eine Aussagenvariable, so gilt (*) nach Definition von $f$.

2. Ist $\varphi = \neg\psi$, so

$$\begin{aligned} f \models \varphi &\iff f \not\models \psi \\ &\iff \mathsf{T}_V \nvdash \psi && \text{nach Ind.vor. für } \psi \\ &\iff \mathsf{T}_V \vdash \neg\psi && \text{da } \mathsf{T}_V \text{ vollständig} \\ &\iff \mathsf{T}_V \vdash \varphi \end{aligned}$$

3. Ist $\varphi = \psi \vee \delta$, so

$$\begin{aligned} f \models \varphi &\iff f \models \psi \textit{ oder } f \models \delta \\ &\iff \mathsf{T}_V \vdash \psi \textit{ oder } \mathsf{T}_V \vdash \delta && \text{nach Ind.vor. für } \psi, \delta \\ &\iff \mathsf{T}_V \vdash \varphi, \end{aligned}$$

wobei sich die letzte Äquivalenz wie folgt ergibt:

Gilt $\mathsf{T}_V \vdash \psi$ oder $\mathsf{T}_V \vdash \delta$, so auch $\mathsf{T}_V \vdash \varphi$ nach der (verallgemeinerten) Expansionsregel.

Sei umgekehrt $\mathsf{T}_V \vdash \varphi$ vorausgesetzt. Falls weder $\mathsf{T}_V \vdash \psi$ noch $\mathsf{T}_V \vdash \delta$, so gilt wegen der Vollständigkeit von $\mathsf{T}_V$: $\mathsf{T}_V \vdash \neg\psi$ und $\mathsf{T}_V \vdash \neg\delta$, also nach 1.6.2 (iii): $\mathsf{T}_V \vdash \neg(\psi \vee \delta)$, d. h. $\mathsf{T}_V \vdash \neg\varphi$, und $\mathsf{T}_V$ wäre damit inkonsistent, Widerspruch! $\square$

Mit Hilfe des Endlichkeitssatzes (aus 1.5.2) erhalten wir als Korollar den

### 1.6.9 Kompaktheitssatz der Aussagenlogik

$$\mathsf{T} \textit{ ist erfüllbar} \iff \textit{jedes endliche } \mathsf{T}_0 \subseteq \mathsf{T} \textit{ ist erfüllbar.}$$

# Kapitel 2

# Strukturen und formale Sprachen

Im Mittelpunkt der modernen Mathematik steht der Begriff einer *Struktur*. Diese kommen vor als

- *relationale Strukturen:*
  Ordnungen, Graphen, Bäume ...
- *algebraische Strukturen:*
  Gruppen, Ringe, Körper, Vektorräume ...
- *gemischte Strukturen:*
  topologische Gruppen, angeordnete Körper, geometrische Räume ...

Verschiedene Strukturen können sich in bestimmten Eigenschaften gleichen oder auch unterscheiden, präzisieren werden wir den Eigenschaftsbegriff mit Hilfe *formaler Sprachen*.

Beliebige Eigenschaften kann man zusammenfassen als Axiome einer (möglicherweise widerspruchsvollen) *Theorie* $T$. Erfüllt eine Struktur $\mathcal{M}$ die Axiome einer Theorie $T$, sind also alle Axiome *wahr* in der Struktur, so nennt man $\mathcal{M}$ ein *Modell* dieser Theorie. Schließlich legen wir fest: Eine Aussage ist (logisch) *wahr* genau dann, wenn sie in *allen* Strukturen wahr ist.

Formale Sprachen haben wir bereits in der Aussagenlogik kennengelernt; wir werden sie jetzt durch zusätzliche Symbole erweitern zu den Sprachen der *Prädikatenlogik*, die dadurch wesentlich ausdrucksstärker werden. Dabei könnten wir ähnlich wie beim Aufbau der Aussagenlogik vorgehen, wollen hier aber zur Abwechslung mit einem semantischen Grundbegriff beginnen. Dafür müssen wir nun allerdings die Kenntnis einiger einfacher Begriffe und Notationen aus der Mengenlehre voraussetzen, die später genauer begründet werden sollen.

## 2.1 Strukturen

Eine **Struktur** $\mathcal{A}$ ist ein 4-Tupel

$$(A,(R_i^{\mathcal{A}}|i\in I),(f_j^{\mathcal{A}}|j\in J),(c_k^{\mathcal{A}}|k\in K)),$$

wobei für jedes $i\in I, j\in J, k\in K$ :

1. $A\neq\emptyset$ eine nicht-leere Menge ist, der **Individuenbereich (Grundbereich, Träger)** der Struktur $\mathcal{A}$,
2. $R_i^{\mathcal{A}}\subseteq A^{n_i}$ ($R_i^{\mathcal{A}}$ ist eine $n_i$-stellige *Relation* auf $A$),
3. $f_j^{\mathcal{A}}:A^{m_j}\longrightarrow A$ ($f_j^{\mathcal{A}}$ ist eine $m_j$-stellige *Funktion* auf $A$),
4. $c_k^{\mathcal{A}}\in A$ ($c_k^{\mathcal{A}}$ ist ein Element (eine *Konstante*) von $A$).

Die Anzahl der Relationen, Funktionen und Konstanten sowie ihre jeweiligen Stellenzahlen (die $\geq 1$ sein sollen) werden im Tripel

$$((n_i|i\in I),(m_j|j\in J),K)$$

festgelegt, genannt **Signatur** oder **Typ** der Struktur. Die Indexmengen $I,J,K$ werden meistens endlich sein, so daß man die einzelnen Relationen, Funktionen und Konstanten auch einfacher hintereinander schreibt (s. folgendes Beispiel). Den Grundbereich einer Struktur $\mathcal{A}$ bezeichnet man auch mit $A=|\mathcal{A}|$, ebenso werden wir den Grundbereich der Struktur $\mathcal{B}$ mit $B$ bezeichnen, usw.

### 2.1.1 Beispiele

1. Die geordneten Gruppen $(\mathbb{Z},<,+,0)$ und $(\mathbb{Q}\setminus\{0\},<,\cdot,1)$ haben eine 2-stellige Relation, eine 2-stellige Funktion und eine Konstante, sind also Strukturen vom Typ ((2),(2),$\{0\}$), aber nicht alle Strukturen dieses Typs müssen geordnete Gruppen sein; man erhält auch eine Struktur $(A,R,f,c)$ dieses Typs durch die Festlegung

$$A:=\mathbb{N},\qquad R:=\{(n,m)\mid n,m\in\mathbb{N},\ n\ teilt\ m+3\},$$
$$f:=\{(n,m,k)\mid n,m,k\in\mathbb{N},\ k=m\cdot n^3+1001\},\quad c:=17,$$

denn die Gültigkeit von Axiomen haben wir (noch) nicht verlangt!

2. Vektorräume $\mathcal{V}$ über einem Körper $K$ fallen unter den obigen Strukturbegriff, wenn man als Grundbereich die Menge $V$ der Vektoren wählt, auf welcher die Vektoraddition als 2-stellige Operation definiert ist. Der Nullvektor als neutrales Element ist eine Konstante von $\mathcal{V}$, und die skalare Multiplikation wird wiedergegeben durch 1-stellige Funktionen $f_\alpha : V \to V$ mit $f_\alpha(v) = \alpha v$ für alle $\alpha \in K$. Im Falle $K = \mathbb{R}$ oder $\mathbb{C}$ ist dann die Indexmenge $J$ überabzählbar (und zwar von der Mächtigkeit des Kontinuums)!

### 2.1.2 Mehrsortige Strukturen

Die Euklidische Ebene ist ein typisches Beispiel einer **mehrsortigen Struktur**: In diesem Fall gibt es zusätzlich eine Menge $S$ von *Sorten*, und statt *eines* Grundbereiches $A$ gibt es nun für jede Sorte $s \in S$ einen Grundbereich $A_s$, Relationen zwischen Elementen möglicherweise verschiedener Sorten, Funktionen, die auf einzelnen Grundbereichen definiert sind und möglicherweise in andere Grundbereiche abbilden und schließlich Konstanten aus den einzelnen Grundbereichen. So gibt es im Falle der Euklidischen Ebene zwei Sorten (Menge der Punkte $P$ bzw. der Geraden $G$) und die Inzidenzbeziehung $I$ zwischen Punkten und Geraden ($p$ *liegt auf* $g$).

Alle Ergebnisse, die hier für einsortige Strukturen dargestellt werden, lassen sich (mutatis mutandis) auf mehrsortige Strukturen übertragen; die mehrsortigen Strukturen zu behandeln, erfordert allerdings einen erheblichen formalen Aufwand wegen der zusätzlichen Sortenindizes. Andererseits lassen sich prinzipiell alle mehrsortigen Strukturen auf einsortige zurückführen, indem man die einzelnen Grundbereiche $A_s$ zu einer disjunkten Vereinigung $A$ zusammenfaßt, für die Aufteilung in die einzelnen Grundbereiche zusätzliche 1-stellige Prädikate $P_s$ auf $A$ einführt und die früheren Relationen und Funktionen in trivialer Weise auf den Grundbereich $A$ fortsetzt. So würde z. B. aus der 2-sortigen Euklidischen Ebene die 1-sortige Struktur

$$\mathcal{E} = (P \cup G, P, G, I) \text{ mit } I = \{(p,g) \mid p \in P, g \in G \textit{ und } p \textit{ liegt auf } g\}$$

Ein Vektorraum $\mathcal{V} = (V, K, +, \cdot, 0)$ über einem Körper $K$ (mit Vektoraddition, Nullvektor und der skalaren Multiplikation $\cdot$ als Abbildung $K \times V \to V$) ist ebenfalls eigentlich eine 2-sortige Struktur (mit Vektoren und Skalaren als den beiden Sorten). Wie wir später sehen werden, ist es aber vorteilhaft, den Vektorraum wie im Beispiel (2) von 2.1.1 als gewöhnliche 1-sortige Struktur aufzufassen, auch wenn damit die Signatur so groß wie der Körper wird.

### 2.1.3 Symbole für formale Sprachen

Die Strukturen $(\mathbb{Z},<,+,0)$ und $(\mathbb{Q}\setminus\{0\},<,\cdot,1)$ haben verschiedene Grundbereiche sowie unterschiedliche Operationen, erfüllen aber beide die Gesetze einer angeordneten Gruppe. Gibt es auch Eigenschaften, in denen sie sich unterscheiden? Um Strukturen desselben Typs hinsichtlich ihrer Eigenschaften untersuchen und vergleichen zu können, müssen wir die Eigenschaften in einer einheitlichen Sprache ausdrücken. So besagt zum Beispiel die Aussage

$$\forall x(e < x \rightarrow \exists y(e < y \wedge y \circ y < x))$$

daß es für jedes Element $x$, welches größer als das neutrale Element ist, ein Element $y$ gibt, welches ebenfalls größer als das neutralen Element ist, so daß das Produkt von $y$ mit $y$ noch kleiner als $x$ ist. Diese Aussage ist offensichtlich in einigen, aber nicht allen Strukturen wahr (das hängt von der *Interpretation* der Symbole ab).

**Definition**

Gegeben sei eine Signatur $\sigma = ((n_i|i \in I),(m_j|j \in J),K)$. Eine **(formale) Sprache** $\mathcal{L}$ zur Signatur $\sigma$ besitzt als **Alphabet** folgende

1. **logische Zeichen**:
   - Variable $v_0, v_1, v_2, \ldots$ (wir schreiben auch $x, x_0, x_1, x_2, y, z, \ldots$)
   - aussagenlogische Verknüpfungen $\neg, \vee$
   - den Existenz-Quantor $\exists$
   - das Gleichheitszeichen $=$ (als 2-stelliges Relationszeichen)

   und dazu - festgelegt durch die Signatur -:

2. **nicht-logische Zeichen**:
   - $n_i$ - stellige Relationszeichen $R_i^{n_i}$ $(n_i \geq 1, i \in I)$,
   - $m_j$ - stellige Funktionszeichen $f_j^{m_j}$ $(m_j \geq 1, j \in J)$,
   - Individuenkonstanten $c_k$ $(k \in K)$.

Sprachen und Strukturen sind somit über die Signatur verbunden; haben die Struktur $\mathcal{A}$ und die Sprache $\mathcal{L}$ dieselbe Signatur, so sagen wir auch, daß $\mathcal{A}$ eine $\mathcal{L}$-**Struktur** ist, bzw. daß $\mathcal{L}$ die **Sprache der Struktur** $\mathcal{A}$ ist.

Als Symbole für die Gruppenoperation bzw. das neutrale Element wählt man üblicherweise $\circ$ bzw. $e$; in konkreten Gruppen, wie

$$\mathcal{A} = (\mathbb{Z}, +, 0) \quad \text{und} \quad \mathcal{B} = (\mathbb{Q}\setminus\{0\}, \cdot, 1),$$

stehen dann die entsprechenden Interpretationen, also in unserer Bezeichnungsweise:

$$\mathcal{A} = (\mathbb{Z}, \circ^{\mathcal{A}}, e^{\mathcal{A}}) \quad \text{und} \quad \mathcal{B} = (\mathbb{Q}\setminus\{0\}, \circ^{\mathcal{B}}, e^{\mathcal{B}}).$$

**Syntax und Semantik**

Die formalen Zeichen (*Symbole*) einer Sprache, ihr Aufbau zu endlichen Zeichenreihen (als *Terme* und *Formeln*) und die Untersuchung ihrer Eigenschaften, die allein auf ihrer formalen Gestalt beruhen, gehören zur **Syntax**. Dagegen handelt es sich beim Begriff der Struktur, sowie allgemeiner den *Interpretationen* der formalen Symbole, um Gegenstände der **Semantik**. Zwischen beiden zu unterscheiden ist besonders wichtig für Grenzbereiche der Mathematischen Logik (GÖDELsche Unvollständigkeitssätze), sollte aber trotzdem schon frühzeitig beachtet werden. Manchmal kann allerdings eine konsequente Wahl der Bezeichnung zu schwerfälligen Ausdrucksweisen führen, so daß wir Vereinfachungen treffen werden. (Z. B. kann das Zeichen $+$ wie oben die Addition auf den ganzen Zahlen bezeichnen, gelegentlich aber auch die Addition auf den natürlichen bzw. reellen Zahlen, möglicherweise aber auch ein formales Symbol für eine 2-stellige Relation - die korrekte Bedeutung wird sich dann aber meistens aus dem Zusammenhang ergeben.)

## 2.2 Unterstrukturen und Morphismen

$\mathcal{A}$ und $\mathcal{B}$ seien Strukturen derselben Sprache $\mathcal{L}$ mit Grundbereich $A$ bzw. $B$. Dann heißt $\mathcal{A}$ **Unterstruktur** (oder **Substruktur)** von $\mathcal{B}$, geschrieben: $\boxed{\mathcal{A} \subseteq \mathcal{B}}$, genau dann, wenn

U1 $A \subseteq B$,

U2 $R^{\mathcal{A}} = R^{\mathcal{B}} \cap A^n$ für jedes $n$-stellige Relationszeichen $R$ von $\mathcal{L}$,

U3 $f^{\mathcal{A}} = f^{\mathcal{B}} \upharpoonright A^m$ für jedes $m$-stellige Funktionszeichen $f$ von $\mathcal{L}$, und

U4 $c^{\mathcal{A}} = c^{\mathcal{B}}$ für jede Individuenkonstante $c$ von $\mathcal{L}$.

Eine Abbildung $h$ heißt **Homomorphismus** von $\mathcal{A}$ nach $\mathcal{B}$, $\boxed{h : \mathcal{A} \longrightarrow \mathcal{B}}$, genau dann, wenn

H1 $h : A \longrightarrow B$,

H2 $R^{\mathcal{A}}(a_1, \ldots, a_n) \Longrightarrow R^{\mathcal{B}}(h(a_1), \ldots, h(a_n))$ für alle $a_1, \ldots, a_n \in A$ und für jedes $n$-stellige Relationszeichen $R$ von $\mathcal{L}$,

H3 $h(f^{\mathcal{A}}(a_1, \ldots, a_m)) = f^{\mathcal{B}}(h(a_1), \ldots, h(a_m))$ für alle $a_1, \ldots, a_m \in A$ und für jedes $m$-stellige Funktionszeichen $f$ von $\mathcal{L}$, und

H4 $h(c^{\mathcal{A}}) = c^{\mathcal{B}}$ für jede Individuenkonstante $c$ von $\mathcal{L}$.

Ein Homomorphismus $h : \mathcal{A} \longrightarrow \mathcal{B}$ heißt eine **Einbettung** gdw $h$ injektiv ist und in (H2) auch die Umkehrung gilt:

H2' $R^{\mathcal{A}}(a_1, \ldots, a_n) \iff R^{\mathcal{B}}(h(a_1), \ldots, h(a_n))$ für alle $a_1, \ldots, a_n \in A$ und für jedes $n$-stellige Relationszeichen $R$ von $\mathcal{L}$.

(Die Injektivität von $h$ besagt gerade, daß (H2') auch für die Gleichheitsrelation gilt.) Ein

- surjektiver Homomorphismus ist ein **Epimorphismus**,
- ein injektiver Homomorphismus ist ein **Monomorphismus**,

  und schließlich ist
- ein **Isomorphismus** eine surjektive Einbettung, also eine bijektive Abbildung $h : A \longleftrightarrow B$, welche die Bedingungen (H1), (H2'), (H3) und (H4) erfüllt.

$\mathcal{A}$ ist **isomorph** zu $\mathcal{B}$, $\boxed{\mathcal{A} \cong \mathcal{B}}$, gdw es einen Isomorphismus $h : \mathcal{A} \longleftrightarrow \mathcal{B}$ gibt.

Somit ist $\mathcal{A} \subseteq \mathcal{B}$ gdw die identische Abbildung $i_A : A \to B$ (mit $i(a) = a$ für alle $a \in A$) eine Einbettung $i_A : \mathcal{A} \longrightarrow \mathcal{B}$ ist.

### 2.2.1 Beispiele

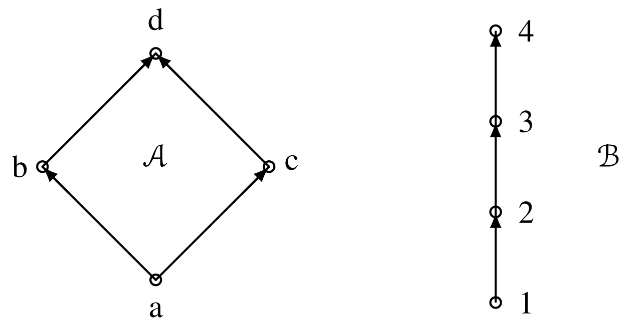
1. Es sei

   - $\mathcal{Z} = (\mathbb{Z}, +, 0)$ die additive Gruppe der ganzen Zahlen,
   - $\mathcal{N} = (\mathbb{N}, +, 0)$ die Struktur der natürlichen Zahlen, und
   - $\mathcal{G} = (\{2n \mid n \in \mathbb{Z}\}, +, 0)$ die Struktur der geraden Zahlen.

   Dann sind $\mathcal{N}, \mathcal{G} \subseteq \mathcal{Z}$, aber $(\mathbb{N}, \cdot, 1) \not\subseteq \mathcal{Z}$. Im übrigen ist $\mathcal{N}$ keine Gruppe: Unterstrukturen von Gruppen brauchen also keine Untergruppen zu sein, da diese auch unter der Inversenbildung abgeschlossen sein müssen. Fügen wir aber zur Sprache ein entsprechendes 1-stelliges Funktionszeichen für das Inverse hinzu, so sind Unterstrukturen zugleich Untergruppen. Ferner ist die Abbildung $h : \mathcal{Z} \longrightarrow \mathcal{G}$ mit $h(n) = 2n$ ein Isomorphismus von $\mathcal{Z}$ auf $\mathcal{G}$.

2. Es sei

   - $\mathcal{A} = (\{a, b, c, d\}, <^A)$ die teilweise geordnete Struktur mit: $a <^A b, b <^A d, a <^A c, c <^A d$,
   - $\mathcal{B} = (\{1, 2, 3, 4\}, <^B)$ die geordnete Struktur mit $1 <^B 2 <^B 3 <^B 4$ :




   Dann ist die Abbildung $h$ mit

   - $h(a) = 1, h(b) = 2, h(c) = 3, h(d) = 4$

   ein Homomorphismus $h : \mathcal{A} \longrightarrow \mathcal{B}$. Da aber $h(b) = 2 <^B h(c) = 3$, während $b \not<^A c$, ist $h$ keine Einbettung, also trotz Bijektivität kein Isomorphismus!

3. Unterstrukturen von Vektorräumen sind Unterräume (lineare Teilräume), wenn man die Sprache der Vektorräume über einem Körper $K$ wie in 2.1.1 mit Hilfe von einstelligen Funktionszeichen $f_\alpha$ für jedes $\alpha \in K$ formuliert, wobei $f_\alpha$ für die skalare Operation mit $\alpha$ steht.

Ist $B$ Universum einer $\mathcal{L}$-Struktur $\mathcal{B}$ und enthält die Sprache $\mathcal{L}$ nur Relationszeichen, so führt *jede* nicht-leere Teilmenge $A \subseteq B$ wieder zu einer Unterstruktur von $\mathcal{B}$: Man schränke einfach die Relationen auf den Teilbereich $A$ ein. Falls die Sprache jedoch auch Funktionszeichen enthält, muß $A$ unter den entsprechenden Funktionen von $\mathcal{B}$ abgeschlossen sein:

### 2.2.2 Satz über Unterstrukturen

*$\mathcal{B}$ sei eine $\mathcal{L}$-Struktur, $\emptyset \neq A \subseteq B$. Dann ist $A$ das Universum einer Unterstruktur $\mathcal{A} \subseteq \mathcal{B}$ genau dann, wenn*

(i) $a_1, \ldots, a_m \in A \Longrightarrow f^{\mathcal{B}}(a_1, \ldots, a_m) \in A$
*für alle m-stelligen Funktionszeichen f von $\mathcal{L}$ und*

(ii) $c^{\mathcal{B}} \in A$ *für jede Individuenkonstante c* von $\mathcal{L}$.

*Sind die obigen Bedingungen erfüllt, so ist die Unterstruktur $\mathcal{A}$ durch die Menge $A$ (und natürlich die Struktur $\mathcal{B}$) eindeutig bestimmt.*

Ist $\mathcal{B}$ eine $\mathcal{L}$-Struktur, $\emptyset \neq A \subseteq B$, so gibt es eine kleinste (bzgl. $\subseteq$) Menge $\overline{A}$, die $A$ enthält und Grundbereich einer Unterstruktur von $\mathcal{B}$ ist: die von $A$ **erzeugte Unterstruktur**. Auf eine Beschreibung von $\overline{A}$ kommen wir nach Einführung des Termbegriffes zurück.

### 2.2.3 Homomorphes Bild, Identifizierung

Ist $h : \mathcal{A} \longrightarrow \mathcal{B}$ ein Homomorphismus, so erfüllt die Menge der Bilder von $h$

$$h[A] := \{h(a) \mid a \in A\}$$

die Bedingungen von Satz 2.2.2, ist also Universum einer Unterstruktur $h[\mathcal{A}]$ von $\mathcal{B}$, das **homomorphe Bild** von $\mathcal{A}$.

Ist $h$ sogar eine Einbettung, so ist $h$ ein Isomorphismus von $\mathcal{A}$ auf $h[\mathcal{A}]$. Häufig **identifiziert** man in dieser Situation $\mathcal{A}$ mit dem isomorphen Bild $h[\mathcal{A}]$, d. h. man ersetzt in $\mathcal{B}$ die Elemente $h(a)$ durch ihre Urbilder $a$ für $a \in A$, um somit $\mathcal{A} \subseteq \mathcal{B}$ zu erhalten (während eigentlich statt $\mathcal{A}$ nur eine isomorphe Kopie hiervon Unterstruktur von $\mathcal{B}$ ist).

## 2.3 Terme einer Sprache

(T1) Jede Variable und jede Individuenkonstante von $\mathcal{L}$ ist ein Term von $\mathcal{L}$.

(T2) Ist $f$ ein $m$-stelliges Funktionszeichen von $\mathcal{L}$ und sind $t_1,\ldots,t_m$ Terme von $\mathcal{L}$, so ist auch $f(t_1,\ldots,t_m)$ ein Term von $\mathcal{L}$.

(T3) Das sind alle Terme von $\mathcal{L}$.

Wie üblich werden wir im Falle 2-stelliger Funktionszeichen (wie + ) statt $f(t,s)$ auch $tfs$ schreiben.

**Beispiele**

- Enthält die Sprache $\mathcal{L}$ keine Funktionszeichen und keine Individuenkonstanten, so sind die Variablen die einzigen $\mathcal{L}$-Terme.
- Im allgemeinen Fall entstehen Terme aus Variablen und Konstanten durch wiederholte Anwendung von Funktionszeichen, z. B.

$$\sin(x^2+z)+e^{i\pi\frac{1}{z}},$$

  was eigentlich schon die Interpretation eines Termes ist, denn hier müßten statt der Funktionen natürlich die entsprechenden Funktionszeichen stehen.

Die *Interpretation* eines Terms in einer Struktur $\mathcal{A}$ wird ein Element von $A$ sein, abhängig davon, wie wir die Variablen im Term belegen (der Wert von $x+y$ hängt von der Operation $+$ sowie von den Werten von $x$ und $y$ ab).

Mit $t(x_1,\ldots,x_n)$ bezeichnen wir einen Term $t$, der (höchstens) die Variablen $x_1,\ldots,x_n$ enthält (möglicherweise weniger - diese Festlegung ist zweckmäßig, da beim Übergang von einem Term zu seinen Teiltermen Variablen verschwinden können). Ein **konstanter Term** ist ein Term ohne Variable.

### 2.3.1 Interpretation von Termen

$\mathcal{A}$ sei eine $\mathcal{L}$-Struktur, $A=|\mathcal{A}|, t=t(x_1,\ldots,x_n)$ ein Term der Sprache $\mathcal{L}$. Ferner seien $a_1,\ldots,a_n$ Elemente von $A$. Dann ist

$t^{\mathcal{A}}[a_1,\ldots,a_n]$ (die **Interpretation** von $t$ in $\mathcal{A}$ unter der Belegung $a_1,\ldots,a_n$)

wie folgt definiert:

$$t^{\mathcal{A}}[a_1,\dots,a_n] = a_i \quad \text{für} \quad t = x_i,$$
$$t^{\mathcal{A}}[a_1,\dots,a_n] = c_k^{\mathcal{A}} \quad \text{für} \quad t = c_k,$$
$$t^{\mathcal{A}}[a_1,\dots,a_n] = f^{\mathcal{A}}(t_1^{\mathcal{A}}[a_1,\dots,a_n],\dots,t_m^{\mathcal{A}}[a_1,\dots,a_n]) \quad \text{für} \quad t = f(t_1,\dots,t_m),$$

wobei $f$ ein $m$-stelliges Funktionszeichen der Sprache $\mathcal{L}$ sei.

Man erhält also $t^{\mathcal{A}}[a_1,\dots,a_n]$, indem man im Term $t$ die Variable $x_i$ durch das Element $a_i$ sowie die Individuenkonstanten und Funktionszeichen in $t$ durch ihre jeweilige Interpretation in $\mathcal{A}$ ersetzt und das Ergebnis in der Struktur $\mathcal{A}$ “auswertet”.

In einem Term selbst kann man für die Variablen keine Objekte direkt einsetzen, da Terme lediglich formale Zeichenreihen sind. Die in der Interpretation vorkommende Belegung ist auch nur eine Zuordnung (daher die Bezeichnung oben mit eckigen Klammern). Wollen wir aber diese Zuordnung in der Sprache selbst ausdrücken, so ist das nur möglich mit Hilfe von Symbolen, welche Namen für die jeweiligen Objekte darstellen, wovon man in der Modelltheorie häufig Gebrauch macht:

### 2.3.2 Spracherweiterung durch Namen

Die Sprache $\mathcal{L}$ wird zur Sprache $\mathcal{L}_A$ erweitert durch Hinzufügen von neuen Individuenkonstanten $\underline{a}$ für jedes Element von $a \in A$; das Symbol $\underline{a}$ wird als **Name** von $a$ bezeichnet.

#### Variante für die Interpretation eines Termes

$t$ sei konstanter Term von $\mathcal{L}_A$ (d. h. $t$ enthalte keine Variablen, u. U. aber Namen für Elemente von $A$). Wir definieren

$$t^{\mathcal{A}} \quad (\text{die } \textbf{Interpretation} \text{ von } t \text{ in } \mathcal{A})$$

durch Induktion über den Aufbau eines Termes wie folgt:

$$\begin{aligned} \underline{a}^{\mathcal{A}} &= a, \\ c_k^{\mathcal{A}} &= c_k^{\mathcal{A}}, \\ f(t_1,\dots,t_m)^{\mathcal{A}} &= f^{\mathcal{A}}(t_1^{\mathcal{A}},\dots,t_m^{\mathcal{A}}), \end{aligned}$$

wobei $f$ ein $m$-stelliges Funktionszeichen $f$ der Sprache $\mathcal{L}$ ist.

Zwischen beiden Arten der Definition besteht offenbar folgender Zusammenhang:

$$t^{\mathcal{A}}[a_1,\ldots,a_n] = (t[\underline{a_1}/x_1,\ldots,\underline{a_n}/x_n])^{\mathcal{A}}.$$

Dabei ist mit $t[\underline{a_1}/x_1,\ldots,\underline{a_n}/x_n]$ auf der rechten Seite der Term gemeint, der aus $t$ entsteht, indem man für $i = 1,\ldots,n$ in $t$ den Term $\underline{a_i}$ für die Variable $x_i$ einsetzt.

Ist $\mathcal{B}$ eine $\mathcal{L}$-Struktur mit $A \subseteq B$, so bezeichnen wir mit $\mathcal{B}_A$ die **Expansion** zu einer $\mathcal{L}_A$-Struktur, die alle Symbole von $\mathcal{L}$ wie in $\mathcal{B}$ interpretiert, jedoch die zusätzlichen neuen Konstanten $\underline{a}$ mit $a$. Statt $\mathcal{A}_A$ schreiben wir häufig kurz $\mathcal{A}$.

Beachte den Unterschied zwischen einer *Erweiterung* und einer *Expansion*:

- $\mathcal{B}$ Erweiterung von $\mathcal{A}$ ($\mathcal{B} \supseteq \mathcal{A}$) : $A$ wird erweitert zu $B$, die Sprache bleibt unverändert,
- $\mathcal{B}$ Expansion von $\mathcal{A}$: die Grundbereiche bleiben gleich ($A = B$), die Sprache wird erweitert (das gilt auch für den allgemeinen Begriff einer Expansion).

Wir können nun den Begriff der erzeugten Unterstruktur neu fassen:

### 2.3.3 Erzeugte Unterstrukturen

*$\mathcal{B}$ sei eine $\mathcal{L}$-Struktur, $A \subseteq B$. Ferner sei $A \neq \emptyset$ oder die Sprache enthalte wenigstens eine Individuenkonstante. Dann gibt es eine kleinste (bzgl. $\subseteq$) Menge $\overline{A}$, die $A$ enthält und Universum einer Unterstruktur von $\mathcal{B}$ ist: die von $A$* **erzeugte Unterstruktur**.

*Man erhält $\overline{A}$, indem man zu $A$ die $c^{\mathcal{B}}$ für jede Individuenkonstante $c$ von $\mathcal{L}$ hinzunimmt und unter den Funktionen $f^{\mathcal{B}}$ abschließt. Alternativ gilt folgende Darstellung:*

$$\begin{aligned}\overline{A} &= \{t^{\mathcal{B}}[a_1,\ldots,a_m] \mid t(v_1,\ldots,v_m)\ \mathcal{L}\text{-Term}, a_1,\ldots,a_m \in A\}\\ &= \{t^{\mathcal{B}_A} \mid t \text{ konstanter } \mathcal{L}\text{-Term von } \mathcal{L}_A\}.\end{aligned}$$

Man kann leicht sehen, daß für Sprachen mit abzählbar-vielen Symbolen und für unendliches $A$ die Erweiterung zu $\overline{A}$ die Mächtigkeit nicht ändert (genauere Aussagen werden wird später der Kardinalzahlbegriffs der Mengenlehre ermöglichen, s. 8.5).

**Beispiele**

1. $(\mathbb{Z}, \leq)$ wird als geordnete Struktur nur von $\mathbb{Z}$ erzeugt, dagegen wird
2. $(\mathbb{Z}, +, -, \cdot, 0, 1)$ als Ring bereits von der leeren Menge erzeugt, während
3. $(\mathbb{Q}, +, -, \cdot, 0, 1)$ von keiner endlichen Menge erzeugt wird.
4. In Vektorräumen (als einsortige Strukturen wie in 2.1.1) sind die erzeugenden Mengen genau die Mengen von Vektoren, deren endliche Linearkombinationen alle Vektoren ergeben (wie im Sinne der *Linearen Algebra*).

## 2.4 Formeln einer Sprache

Terme bezeichnen *Objekte* einer Struktur, Formeln *Eigenschaften*. Ähnlich wie der Begriff eines Terms werden Formeln einer Sprache $\mathfrak{L}$ ($\mathfrak{L}$-**Formeln**) durch *Rekursion* definiert:

(F1) Die einfachsten $\mathfrak{L}$-Formeln (auch **atomare** oder **Prim**-Formeln genannt) sind:

- $s = t$ für beliebige Terme $s, t$ und
- $R(t_1, \ldots, t_n)$ für jedes $n$-stellige Relationszeichen von $\mathfrak{L}$ und beliebige $\mathfrak{L}$-Terme $t_1, \ldots, t_n$.

(F2) Ist $\varphi$ eine $\mathfrak{L}$-Formel, so auch $\neg\varphi$.

(F3) Sind $\varphi$ und $\psi$ $\mathfrak{L}$-Formeln, so auch $(\varphi \vee \psi)$.

(F4) Ist $\varphi$ eine $\mathfrak{L}$-Formel und $x$ eine Variable, so ist auch $\exists x\, \varphi$ eine $\mathfrak{L}$-Formel.

(F5) Das sind alle $\mathfrak{L}$-Formeln.

Wiederum werden wir im Falle 2-stelliger Relationszeichen (wie $=$ oder $<$ ) statt $R(t, s)$ auch $tRs$ schreiben; während wir jedoch im ersten Fall (wo die Relationszeichen stets vor den Argumenten stehen) auf Klammern verzichten können, müssen wir im zweiten Fall jedoch möglicherweise Klammern setzen, um eindeutige Lesbarkeit zu sichern.

Die aussagenlogischen Zeichen $\wedge, \rightarrow$ und $\leftrightarrow$ sowie den Allquantor $\forall$ definieren wir:

$$\begin{aligned}
\varphi \wedge \psi \quad &:\Longleftrightarrow \quad \neg(\neg\varphi \vee \neg\psi), \\
\varphi \rightarrow \psi \quad &:\Longleftrightarrow \quad \neg\varphi \vee \psi, \\
\varphi \leftrightarrow \psi \quad &:\Longleftrightarrow \quad (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi), \\
\forall x\varphi \quad &:\Longleftrightarrow \quad \neg\exists x\neg\varphi.
\end{aligned}$$

Somit stellt jede (prädikatenlogische) Sprache $\mathcal{L}$ eine Erweiterung der aussagenlogischen Sprache dar, indem die Aussagenvariable der Aussagenlogik durch atomare Formeln ersetzt werden und zusätzlich als logische Symbole die Gleichheit und die Quantoren auftreten. Formel- und Termbegriff sind wieder rekursiv definiert, so daß man Aussagen über alle Formeln (bzw. Terme) wiederum durch Induktion durchzuführen kann:

### 2.4.1 Beweis durch Induktion über den Formelaufbau

*Es sei E eine Eigenschaft mit*

(i) *E($\varphi$) gilt für alle atomaren Formeln $\varphi$,*

(ii) *gilt $E(\varphi)$, so auch $E(\neg\varphi)$,*

(iii) *gelten $E(\varphi)$ und $E(\psi)$, so auch $E(\varphi \vee \psi)$, und*

(iv) *gilt $E(\varphi)$, so auch $E(\exists x\, \varphi)$.*

*Dann gilt $E(\varphi)$ für alle Formeln $\varphi$.*

Um eine Eigenschaft für alle Formeln nachzuweisen, muß man sie also

(1) für alle atomaren Formeln nachweisen und dann

(2) unter der Annahme, daß die Eigenschaft bereits für die Formel $\varphi$ (bzw. die Formeln $\varphi$ und $\psi$) gilt, für die Formeln $\neg\varphi, (\varphi \vee \psi)$ und $\exists x\, \varphi$ nachweisen.

(1) nennt man wieder den **Induktionsanfang**, (2) den **Induktionsschluß**, welcher aus einer **Induktionsannahme** und einer **Induktionsbehauptung** besteht. (Ähnliches gilt für die Beweismethode der Induktion über den Termaufbau und auch das Induktionsverfahren für ein geeignetes Komplexitätsmaß, s. 1.1.3.)

### 2.4.2 Freie und gebundene Variable

Bildet man eine Formel $\exists x\,\varphi$ nach (F4), so heißt die in $\varphi$ vorkommende Variable **gebunden** (durch den Quantor $\exists$); ähnlich natürlich im Falle der Formel $\forall x\,\varphi$. Die übrigen Variablen heißen **freie** Variable, weil sie frei für Einsetzungen sind. Während es nicht sinnvoll ist, in gebundene Variable Einsetzungen vorzunehmen, kann man sie aber in eine andere geeignete Variable umbenennen, ohne die Bedeutung (die wir noch erklären werden) zu ändern.

Bei obiger Definition treten nun aber folgende Probleme auf:

- Bildet man aus zwei Formeln $\varphi, \psi$ nach (F3) die neue Formel $(\varphi \vee \psi)$, so kann es vorkommen, dass dieselbe Variable $y$ in der einen Formel frei, in der anderen jedoch gebunden vorkommt:

$$(*) \qquad \neg x = y \rightarrow \exists y(x < y \wedge 1 < x \circ y)$$

- In (F4) ist nicht ausgeschlossen, daß die Variable $x$ in $\varphi$ bereits (in einer Teilformel) gebunden vorkommt.

Prinzipiell bereiten beide Probleme keine Schwierigkeiten; statt zwischen freien und gebundenen Variablen zu unterscheiden, spricht man vom freien bzw. gebundenen *Vorkommen* einer Variable in einer Formel, so daß dieselbe Variable in einer Formel an einigen Stellen frei, an anderen jedoch auch gebunden vorkommen kann. Lästig ist diese Freiheit der Formelbildung jedoch bei *Einsetzungen*, welche besondere Vorkehrungen erfordern: Will man in obiger Formel (*) einen Term für eine Variable einsetzen, so muß man sich auf Stellen beschränken, an denen die Variable frei vorkommt. Wichtiger ist jedoch zu beachten, daß Einsetzungen so erfolgen, daß die Formel

$\varphi(t)$ über den Term $t$ dasselbe aussagt
wie $\varphi(x)$ über die Variable $x$,

und dazu muß man zusätzlich darauf achten, daß beim Einsetzen eines Terms $t(y\ldots)$ in eine frei vorkommende Variable $x$ nicht durch den Term $t$ eine Variable $y$ in den Bindungsbereich eines Quantors gerät, z. B. wenn man $(y+1)$ für $x$ in die Formel (*) einsetzt, in welcher $x$ überall frei vorkommt!

Diese Probleme kann man durch die Wahl von unterschiedlichen Variablen für freie bzw. gebundene Variable umgehen (wie in der älteren HILBERT-Schule). Wir wollen jedoch bei dem heute allgemein üblichen Gebrauch einer einzigen Sorte von Variablen bleiben, andererseits aber Einsetzungen leichter handhaben und daher folgende Vereinbarungen treffen (die man intuitiv zwecks besserer Lesbarkeit ohnehin wohl beachten würde):

- Bei der Bildung einer Formel $\varphi$ sind gebundene Variablen so zu wählen, daß sie nicht zugleich auch frei vorkommen.

Unter dieser Voraussetzung kann man dann wieder einfach von freien bzw. gebundenen Variablen sprechen (anstatt von freien bzw. gebundenen Vorkommen).

Terme enthalten auf jeden Fall nur freie Variable, während in Formeln sowohl freie wie auch gebundene Variable vorkommen können. Da für die Interpretation von Termen bzw. Formeln in einer Struktur Belegungen der freien Variablen angegeben werden müssen, schreiben wir oft

$$t(x_1,\ldots,x_n) \quad \text{bzw.} \quad \varphi(x_1,\ldots,x_n)$$

um anzuzeigen, daß alle im Term $t$ vorkommenden Variable bzw. in der Formel $\varphi$ vorkommenden *freien* Variable unter den $x_1,\ldots,x_n$ vorkommen.

### 2.4.3 Substitution

$\varphi$ sei eine Formel, $s,t$ Terme, $x$ eine freie Variable. Dann entstehe

- $s[t/x]$ aus $s$, indem überall $t$ für $x$ eingesetzt wird, ebenso
- $\varphi[t/x]$ aus $\varphi$, indem überall $t$ für $x$ eingesetzt wird.

Dabei soll im zweiten Fall $t$ **substituierbar** in $\varphi$ sein, d. h. $t$ soll keine Variable enthalten, die in $\varphi$ gebunden vorkommt.

Ähnlich erklären wir die mehrfache Substitution

$$s[t_1/x_1,\ldots,t_n/x_n] \quad \text{und} \quad \varphi[t_1/x_1,\ldots,t_n/x_n]$$

durch mehrfache (simultane) Ausführung der jeweiligen Substitution.

Wenn man aus dem Zusammenhang weiß, für welche Variable eine Substitution erfolgt (oder wenn man einfach bequemer sein will), schreibt man kürzer auch $s(t)$ für $s[t/x]$ und $\varphi(t)$ für $\varphi[t/x]$, und ähnlich im Falle mehrfacher Substitution. Natürlich kann man auch hintereinander substituieren: erst $t_1$ für $x_1$, dann $t_2$ für $x_2$, usw.; wenn dann aber mit $t_1$ auch die Variable $x_2$ hineinkommt, so wird auch hier (anders als bei der simultanen Substitution) anschließend der Term $t_2$ eingesetzt!

**Definition**

$$\begin{aligned} t\ \textbf{konstanter}\ \text{Term}: &\iff t\ \text{Term ohne Variable},\\ \varphi\ \textbf{Satz (Aussage)}: &\iff \varphi\ \text{Formel ohne freie Variable}. \end{aligned}$$

Der Begriff *Satz* im Sinne der obigen Definition ist rein syntaktisch und natürlich nicht mit dem Begriff des Satzes im Sinne eines *Theorems* zu verwechseln! Im Englischen unterscheidet man oft zwischen *sentence* und *theorem*; gelegentlich nennt man Sätze im Sinne der obigen Definition auch *Aussagen* oder *abgeschlossene Formeln*, um eine Verwechslung mit der Bedeutung als Theorem zu vermeiden.

# Kapitel 3

# Modelle und Theorien

## 3.1 Das Wahrheitsprädikat: Modelle

Nachdem wir die Interpretation von Termen in einer Struktur definiert haben, werden wir in ähnlicher Weise die *Bedeutung* einer Formel in einer Struktur erklären. Gegeben sei

- eine $\mathcal{L}$-Struktur $\mathcal{A}$ mit Grundbereich $A = |\mathcal{A}|$,
- eine $\mathcal{L}$-Formel $\varphi = \varphi(x_1, \ldots, x_n)$, deren freie Variable unter $x_1, \ldots, x_n$ vorkommen,
- eine **Belegung** $f = (a_1, \ldots, a_n)$ der Variablen $x_1, \ldots, x_n$.

Dann definieren wir induktiv:[1]

$$\boxed{\mathcal{A} \models \varphi[f]} \qquad \varphi \text{ ist } \textbf{wahr} \text{ in } \mathcal{A} \text{ unter der } \textbf{Belegung } f$$

wie folgt:

$$\begin{array}{rcc}
\mathcal{A} \models (t = s)[f] & \iff & t^{\mathcal{A}}[f] = s^{\mathcal{A}}[f] \\
\mathcal{A} \models R(t_1, \ldots, t_n)[f] & \iff & (t_1^{\mathcal{A}}[f], \ldots, t_n^{\mathcal{A}}[f]) \in R^{\mathcal{A}} \\
\mathcal{A} \models \neg\varphi[f] & \iff & \mathcal{A} \not\models \varphi[f] \\
\mathcal{A} \models (\varphi \vee \psi)[f] & \iff & \mathcal{A} \models \varphi[f] \textit{ oder } \mathcal{A} \models \psi[f] \\
\mathcal{A} \models \exists x_i \varphi[f] & \iff & \textit{es gibt ein } a \in A \textit{ mit } \mathcal{A} \models \varphi[f(a/x_i)],
\end{array}$$

wobei $f(a/x_i)$ die Belegung $f$ ist, allerdings mit $a$ als Belegung von $x_i$.

[1] A. Tarski: *Der Wahrheitsbegriff in den formalisierten Sprachen,* Studia Phil. 1 (1936)

Auch hier läßt sich eine alternative Definition angeben, die formal etwas einfacher ausfällt:

**Definition des Wahrheitsprädikats: Variante in erweiterter Sprache $\mathcal{L}_A$**

$\mathcal{A}$ sei eine $\mathcal{L}$-Struktur, $A = |\mathcal{A}|$, $\sigma$ sei ein $\mathcal{L}_A$-Satz. Wir definieren induktiv

$$\mathcal{A} \models \sigma \quad \text{(genauer: } \mathcal{A}_A \models \sigma\text{)}$$

wie folgt:

$$\begin{aligned}
\mathcal{A} \models (t = s) \quad &\iff \quad t^{\mathcal{A}} = s^{\mathcal{A}} \\
\mathcal{A} \models R(t_1, \ldots, t_n) \quad &\iff \quad (t_1^{\mathcal{A}}, \ldots, t_n^{\mathcal{A}}) \in R^{\mathcal{A}} \\
\mathcal{A} \models \neg\sigma \quad &\iff \quad \mathcal{A} \not\models \sigma \\
\mathcal{A} \models (\sigma \vee \tau) \quad &\iff \quad \mathcal{A} \models \sigma \textit{ oder } \mathcal{A} \models \tau \\
\mathcal{A} \models \exists x\varphi \quad &\iff \quad \textit{es gibt ein } a \in A \textit{ mit } \mathcal{A} \models \varphi[\underline{a}/x]
\end{aligned}$$

Wiederum haben wir hier eine einfache Beziehung zwischen diesen beiden Definitionen:

$$\mathcal{A} \models \varphi[a_1, \ldots, a_n] \iff \mathcal{A} \models \varphi[\underline{a_1}/x_1, \ldots, \underline{a_n}/x_n].$$

Dabei steht auf der linken Seite eine $\mathcal{L}$-Formel $\varphi$ mit höchstens $x_1, \ldots, x_n$ als frei vorkommenden Variablen, die mit $a_1, \ldots, a_n$ belegt werden, auf der rechten Seite (bezogen auf die alternative Definition) ein $\mathcal{L}_A$-Satz $\varphi$, welcher an Stelle der $x_1, \ldots, x_n$ Namen für die $a_1, \ldots, a_n$ enthält. (Wenn nicht ausdrücklich darauf hingewiesen wird, welche der beiden Definitionen gebraucht wird, muß dies aus dem Zusammenhang erschlossen werden.)

Für eine $\mathcal{L}$-Formel $\varphi = \varphi(x_1, \ldots, x_n)$ sei $\varphi^{\forall} := \forall x_1 \ldots \forall x_n \varphi$ der **universelle Abschluß** von $\varphi$. Wir definieren dann

$\varphi$ ist **wahr** in $\mathcal{A}$ (oder: $\mathcal{A}$ ist **Modell** von $\varphi$)

durch

$$\boxed{\mathcal{A} \models \varphi :\iff \mathcal{A} \models \forall x_1 \ldots \forall x_n \varphi}$$

Diese Festlegung entspricht dem üblichen Gebrauch in der Mathematik:

Die Axiome einer algebraischen Theorie schreibt man oft als Formeln auf, z. B. $a+b=b+a$ im Falle des Kommutativgesetzes der Gruppentheorie, meint damit aber, daß es in einer Gruppe für alle Elemente gelten soll, d. h. in der Form $\forall x \forall y(x+y=y+x)$.

Die **Wahrheit** (oder **Allgemeingültigkeit**) einer $\mathcal{L}$-Formel $\varphi = \varphi(x_1,\ldots,x_n)$ wird definiert durch

$$\boxed{\models \varphi :\Longleftrightarrow \textit{ für alle } \mathcal{L}\textit{-Strukturen } \mathcal{A} \textit{ gilt: } \mathcal{A} \models \varphi}$$

Das bedeutet also für eine Formel $\varphi$, daß sie bei *allen möglichen* Interpretationen der nicht-logischen Zeichen und für *alle* Einsetzungen in die frei vorkommenden Variablen wahr sein muß. Ebenso definieren wir im Falle einer Menge $\Sigma$ von Formeln: $\mathcal{A}$ ist **Modell** von $\Sigma$ gdw $\mathcal{A}$ ist Modell von allen $\varphi \in \Sigma$:

$$\boxed{\mathcal{A} \models \Sigma :\Longleftrightarrow \textit{ für alle } \varphi \in \Sigma : \mathcal{A} \models \varphi}.$$

Schließlich erklären wir den semantischen Begriff der **Folgerung**: Eine Formel $\varphi$ **folgt** aus der Formelmenge $\Sigma$ gdw sie in allen Modellen von $\Sigma$ gilt:

$$\boxed{\Sigma \models \varphi :\Longleftrightarrow \textit{ für alle } \mathcal{A} \models \Sigma \textit{ gilt: } \mathcal{A} \models \varphi}$$

Endliche Folgen von Variablen $x_0,\ldots,x_n$ bezeichnen wir auch mit $\vec{x}$ (ähnlich in anderen Fällen). Die Bedeutung der Einführung neuer Konstanten zeigt sich in dem folgenden

### 3.1.1 Satz über neue Konstanten

*Es sei $\Gamma$ eine Menge von Sätzen der Sprache $\mathcal{L}$, $\varphi(\vec{x})$ eine $\mathcal{L}$-Formel und $\vec{c}$ eine Folge von Individuenkonstanten (derselben Länge wie $\vec{x}$), die in der Sprache $\mathcal{L}$ nicht vorkommen. Dann gilt:*

$$\Gamma \models \varphi(\vec{c}) \Longrightarrow \Gamma \models \forall \vec{x}\, \varphi(\vec{x}).$$

*Beweis:* Sei $\Gamma \models \varphi(\vec{c})$, $\mathcal{A} \models \Gamma$ und $\vec{a}$ seien $n$ beliebige Elemente von $A$. Wir müssen zeigen, daß $\mathcal{A} \models \varphi[\vec{a}]$. Dazu sei $\mathcal{L}'$ die Sprache $\mathcal{L}$, erweitert um die neuen Individuenkonstanten $\vec{c}$. Wir expandieren $\mathcal{A}$ zu einer $\mathcal{L}'$-Struktur $\mathcal{A}'$, indem wir $c_i^{\mathcal{A}'} = a_i$ setzen. Dann ist auch $\mathcal{A}' \models \Gamma$ (da die neuen Konstanten in der Menge $\Gamma$

nicht vorkommen!), also gilt nach Voraussetzung $\mathcal{A}' \models \varphi(\vec{c})$, d. h. $\mathcal{A} \models \varphi[\vec{a}]$, weil wir $c_i^{\mathcal{A}'} = a_i$ festgelegt haben. $\square$

Dieser Satz bedeutet also, daß eine Aussage der Form $\forall \vec{x}\, \varphi(\vec{x})$ aus einer Menge von Sätzen $\Gamma$ folgt, wenn sie für “feste, aber beliebige” Elemente folgt. Meistens wird man $\Gamma$ als die Menge von Axiomen einer Theorie auffassen:

## 3.2 Axiome und Theorien

Eine **Theorie** $\mathsf{T}$ ist bestimmt durch

1. eine formale Sprache $\mathcal{L}$, die *Sprache* von $\mathsf{T}$, sowie
2. eine Menge $\Sigma$ von Formeln der Sprache $\mathcal{L}$ als *Axiome* der Theorie.

Da man die Sprache einer Theorie meistens aus den Axiomen ablesen kann, werden wir in der Regel eine Theorie mit der Menge $\Sigma$ ihrer Axiome gleichsetzen.

Somit erklären wir für eine Theorie $\mathsf{T}$ mit der Axiomenmenge $\Sigma$ und eine Formel $\varphi$:

$$
\begin{array}{rclr}
\mathcal{A} \models \mathsf{T} :\Longleftrightarrow & & \mathcal{A} \models \Sigma & \mathcal{A} \textbf{ ist Modell von } \mathsf{T}, \\
\mathsf{T} \models \varphi :\Longleftrightarrow & & \textit{für alle } \mathcal{A} \models \mathsf{T}:\ \mathcal{A} \models \varphi & \varphi \textbf{ folgt aus } \mathsf{T}.
\end{array}
$$

$\mathsf{T} + \varphi$ bezeichnet die Theorie $\mathsf{T}$ mit $\varphi$ als zusätzlichem Axiom.

Jeder Theorie $\mathsf{T}$ (oder Menge von Sätzen) kann man die Klasse ihrer Modelle zuordnen:

$$Mod(\mathsf{T}) := \{\mathcal{A} \mid \mathcal{A} \models \mathsf{T}\} \qquad \textbf{die Modellklasse von } \mathsf{T},$$

und umgekehrt jeder $\mathcal{L}$-Struktur $\mathcal{A}$ eine Menge von Sätzen, ihre *Theorie*:

$$\mathsf{Th}(\mathcal{A}) := \{\sigma \mid \sigma\ \mathcal{L}\text{-Satz}, \mathcal{A} \models \sigma\} \quad \textbf{die (elementare) Theorie von } \mathcal{A}.$$

Häufig verlangt man von Theorien, daß sie *deduktiv abgeschlossen* sind, zählt also auch alle Folgerungen zur Theorie. Stattdessen werden wir Theorien als gleich ansehen, wenn sie dieselben Folgerungen haben (also sich höchstens durch die Wahl ihrer Axiome unterscheiden), und diese Aussage bedeutet wiederum, daß sie dieselben Modelle haben:

$$
\begin{aligned}
\mathsf{T} = \mathsf{T}' : \quad &\Longleftrightarrow \textit{für alle Sätze } \sigma : \mathsf{T} \models \sigma \iff \mathsf{T}' \models \sigma \\
&\Longleftrightarrow Mod(\mathsf{T}) = Mod(\mathsf{T}').
\end{aligned}
$$

Der oben eingeführte semantische Folgerungsbegriff hat im konkreten Fall meistens einen schwerwiegenden Nachteil: Denn semantisch folgt etwa eine Aussage $\sigma$ aus der Gruppentheorie, wenn sie in *allen* Gruppen gilt. Das kann man nun aber nicht nachweisen, indem man die Gültigkeit der Aussage in allen einzelnen Gruppen nachprüft, schon die Wahrheit einer Formel in *einer* (unendlichen) Struktur ist i. a. schwierig zu überprüfen.

**Der syntaktische Folgerungsbegriff: Beweise**

Folgerungen aus einer mathematischen Theorie wird man versuchen, mit *logischen Argumenten* zu begründen, indem man nach einem *Beweis* einer Formel $\varphi$ aus den Axiomen der Theorie sucht. Wir werden im nächsten Kapitel einen formalen Begriff eines Beweises einführen, der sich auf ein Axiomensystem der Prädikatenlogik bezieht. Wenn $\mathsf{T} \vdash \varphi$ besagt, daß es einen Beweis von $\varphi$ aus den Axiomen von $\mathsf{T}$ mit Hilfe logischer Axiome und Regeln gibt, so wird für geeignete Axiomensysteme der Logik der **Vollständigkeitssatz** in der folgenden Form gelten:

$$\mathsf{T} \models \varphi \iff \mathsf{T} \vdash \varphi.$$

## 3.3 Der Kompaktheitssatz und einige Folgerungen

Da Beweise endlich sein müssen, kann ein Beweis einer Aussage aus den Axiomen einer Theorie auch nur von endlich-vielen Axiomen Gebrauch machen. Zusammen mit dem oben erwähnten Vollständigkeitssatz folgt daraus sofort der für die Modelltheorie zentrale

### 3.3.1 Kompaktheitssatz

*Für jede Theorie* $\mathsf{T}$ *und jeden Satz* $\sigma$ *der Sprache von* $\mathsf{T}$ *gilt:*

(i) $\mathsf{T} \models \sigma \implies$ *es gibt ein endliches* $\mathsf{T}_0 \subseteq \mathsf{T}$ *mit* $\mathsf{T}_0 \models \sigma$,

(ii) *jedes endliche* $\mathsf{T}_0 \subseteq \mathsf{T}$ *hat ein Modell* $\implies \mathsf{T}$ *hat ein Modell.*

(Trivialerweise gelten natürlich auch die jeweiligen Umkehrungen.)

Der Kompaktheitssatz ist aber ein rein modelltheoretischer Satz und soll später auch mit modelltheoretischen Methoden bewiesen werden. Hier sei schon auf einige Anwendungen hingewiesen:

- Man kann zeigen, daß viele Theorien Modelle besitzen, an die man bei der Axiomatisierung "eigentlich nicht gedacht" hat: Nicht-Standardmodelle der PEANO-Arithmetik, die unendlich-große Zahlen besitzen, oder Nicht-Archimedisch angeordnete Körper.

- Einige Klassen von Modellen lassen sich (in der Sprache der Prädikatenlogik erster Stufe) nicht axiomatisch erfassen.

Um hierfür ein Beispiele zu erhalten, benutzen wir

**Anzahlformeln**

In der Sprache der Prädikatenlogik lassen sich allein mit Hilfe der Gleichheit Sätze formulieren, die Aussagen über die Anzahl der Elemente einer Struktur machen (explizite Angabe als Übungsaufgabe):

$$\begin{aligned} \mathcal{A} \models \exists^{\geq n} &\iff \textit{A hat mindestens n Elemente}, \\ \mathcal{A} \models \exists^{\leq n} &\iff \textit{A hat höchstens n Elemente}, \\ \mathcal{A} \models \exists^{n} &\iff \textit{A hat genau n Elemente}. \end{aligned}$$

Die Menge der Sätze

$$\Sigma_\infty := \{\exists^{\geq n} | n \in \mathbb{N}\}$$

hat also nur *unendliche* Modelle; dagegen kann man i. a. durch keine Menge von Sätzen die *Endlichkeit* eines Modells erzwingen:

### 3.3.2 Satz über die Existenz unendlicher Modelle

$\mathsf{T}$ *sei eine Theorie, die beliebig große endliche Modelle hat (d.h. für jede natürliche Zahl n hat* $\mathsf{T}$ *ein Modell mit mindestens n Elementen). Dann hat* $\mathsf{T}$ *auch ein unendliches Modell.*

*Beweis:* Nimm zur Theorie $\mathsf{T}$ die oben definierte Satzmenge $\Sigma_\infty$ hinzu und nenne die erweiterte Theorie $\mathsf{T}'$. Dann gilt nach Voraussetzung über $\mathsf{T}$:

jede endliche Teilmenge von $\mathsf{T}'$ hat ein Modell,

also hat nach dem Kompaktheitssatz $\mathsf{T}'$ selbst ein Modell, das dann Modell von $\mathsf{T}$ ist und nicht endlich sein kann. $\square$

## 3.4 Diagramme

Die Eigenschaften einer $\mathcal{L}$-Struktur $\mathcal{A}$ werden durch ihre **Theorie** $\mathsf{Th}(\mathcal{A})$ beschrieben. Um auch die Eigenschaften anzugeben, die einzelne Elemente von $\mathcal{A}$ besitzen, gehen wir wieder zur erweiterten Sprache $\mathcal{L}_A$ über. A. ROBINSON hat die (für die Modelltheorie wichtige) Methode der *Diagramme* eingeführt: Für eine Menge $\mathsf{H}$ von $\mathcal{L}$-Formeln ist das $\mathsf{H}$-**Diagramm** von $\mathcal{A}$ die Menge der Sätze

- $\mathsf{D}_\mathsf{H}(\mathcal{A}) := \{\varphi(\underline{a_1},\ldots,\underline{a_n}) \mid \varphi \in \mathsf{H},\ \mathcal{A} \models \varphi[a_1,\ldots,a_n]\}$, speziell
- $\mathsf{D}_+(\mathcal{A}) := \mathsf{D}_{\mathsf{At}}(\mathcal{A})$ das **positive Diagramm** von $\mathcal{A}$
für $\mathsf{H} = \mathsf{At}$ = Menge der atomaren $\mathcal{L}$-Formeln,
- $\mathsf{D}(\mathcal{A}) := \mathsf{D}_{\mathsf{Ba}}(\mathcal{A})$ das quantorenfreie (oder ROBINSON-)**Diagramm** von $\mathcal{A}$, für $\mathsf{H} = \mathsf{Ba}$ = Menge der $\mathcal{L}$-Formeln, welche atomar oder Negation einer atomaren Formel sind (**Basis**-Formeln oder **Literale**),
- $\mathsf{D}_L(\mathcal{A}) = \mathsf{Th}(\mathcal{A}_A)$ das **elementare Diagramm** von $\mathcal{A}$.

Das positive Diagramm kann man sich als eine Verallgemeinerung der Gruppentafel vorstellen; im Falle von $(\mathbb{N},\cdot)$ enthält es das kleine und große Einmaleins: $0\cdot 0 = 0, 0\cdot 1 = 0, \ldots, 1\cdot 1 = 1, \cdots, 2\cdot 3 = 6$, aber auch $2\cdot(5\cdot 3) = 30, (1\cdot 2)\cdot 6 = 3\cdot 4$, usw. (genauer immer mit $\underline{n}$ statt $n$ geschrieben), das ROBINSON-Diagramm enthält zusätzlich auch die Ungleichungen $1\cdot 4 \neq 2, 3\cdot 2 \neq 5, \ldots$ Dagegen ist $\mathsf{Th}(\mathcal{A}_A)$, das elementare Diagramm von $\mathcal{A}$, die umfangreichste Beschreibung der Eigenschaften aller Elemente von $A$, die in der Struktur $\mathcal{A}$ gelten.

Offensichtlich ist $\mathcal{A}_A$ ein Modell von $\mathsf{D}(\mathcal{A})$ und sogar von $\mathsf{Th}(\mathcal{A}_A)$. Es ist leicht zu sehen, daß für $\mathcal{A} \subseteq \mathcal{B}$ auch die Struktur $\mathcal{B}_A$ ein Modell von $\mathsf{D}(\mathcal{A})$ ist; allgemein wird ein Modell von $\mathsf{D}(\mathcal{A})$ eine $L_A$-Struktur der Form $(\mathcal{B}, (h(a))_{a\in A})$ sein, wobei $h(a) = \underline{a}^{\mathcal{B}}$ durchaus von $a$ verschieden sein kann. Jedoch kann man in diesem Fall $\mathcal{A}$ mit einer Substruktur von $\mathcal{B}$ identifizieren:

### 3.4.1 Diagrammlemma

*Es sei $\mathcal{A}$ eine $\mathcal{L}$-Struktur, $(\mathcal{B}, (\underline{a}^{\mathcal{B}})_{a\in A})$ eine $\mathcal{L}_A$-Struktur und $h : A \longrightarrow B$ definiert durch $h(a) = \underline{a}^{\mathcal{B}}$. Dann gilt:*

(i) $(\mathcal{B}, (h(a))_{a\in A}) \models \mathsf{D}(\mathcal{A}) \iff h : \mathcal{A} \longrightarrow \mathcal{B}$ *ist eine Einbettung,*

(ii) $\mathcal{B}_A \models \mathsf{D}(\mathcal{A}) \iff \mathcal{A} \subseteq \mathcal{B}$, *falls* $A \subseteq B$.

*Beweis* von (i): Wir schreiben kurz $\mathcal{B}_h$ für $(\mathcal{B},(h(a))_{a\in A})$ und nehmen an, daß $\mathcal{B}_h \models \mathsf{D}(\mathcal{A})$. Dann gilt für beliebige $a,b \in A$:

$$\begin{aligned} a \neq b &\iff \underline{a} \neq \underline{b} \in \mathsf{D}(\mathcal{A}) \\ &\implies \quad \mathcal{B}_h \models \underline{a} \neq \underline{b} \quad \text{nach Vor.} \\ &\iff h(a) \neq h(b), \end{aligned}$$

also ist $h$ injektiv. Ähnlich im Falle der Relationszeichen:

$$\begin{aligned} R^{\mathcal{A}}(a_1,\dots,a_n) &\iff R(\underline{a_1},\dots,\underline{a_n}) \in \mathsf{D}(\mathcal{A}) \\ &\implies \quad \mathcal{B}_h \models R(\underline{a_1},\dots,\underline{a_n}) \quad \text{nach Vor.} \\ &\iff R^{\mathcal{B}}(h(a_1),\dots,h(a_n)). \end{aligned}$$

Das gleiche Argument für die Formel $\neg R$ liefert die Umkehrung, also

$$R^{\mathcal{A}}(a_1,\dots,a_n) \iff R^{\mathcal{B}}(h(a_1),\dots,h(a_n)),$$

und schließlich gilt auch für die Funktionszeichen:

$$\begin{aligned} f^{\mathcal{A}}(a_1,\dots,a_m) = a &\iff f(\underline{a_1},\dots,\underline{a_n}) = \underline{a} \in \mathsf{D}(\mathcal{A}) \\ &\implies \quad \mathcal{B}_h \models f(\underline{a_1},\dots,\underline{a_n}) = \underline{a} \quad \text{nach Vor.} \\ &\iff f^{\mathcal{B}}(h(a_1),\dots,h(a_n) = h(a). \end{aligned}$$

Mit denselben Argumenten zeigt man auch die umgekehrte Richtung von (i). □

### 3.4.2 Universelle Theorien

$$\begin{aligned} \psi \textit{ ist eine } \forall\textit{-Formel (oder } \textbf{universell}\textit{):} &\iff \psi = \forall x_1 \dots \forall x_n\ \varphi \\ &\quad \textit{für eine quantorenfreie Formel } \varphi, \\ \psi \textit{ ist eine } \exists\textit{-Formel (oder } \textbf{existentiell}\textit{):} &\iff \psi = \exists x_1 \dots \exists x_n\ \varphi \\ &\quad \textit{für eine quantorenfreie Formel } \varphi. \end{aligned}$$

Für eine Theorie $\mathsf{T}$ sei

$$\mathsf{T}_\forall := \{\sigma \mid \mathsf{T} \models \sigma,\ \sigma \textit{ ein } \forall\textit{-Satz}\} \quad \text{der } \textbf{universelle Teil} \text{ von } \mathsf{T}.$$

Eine Theorie $\mathsf{T}$ heiße **Substruktur-abgeschlossen** gdw

$$\mathcal{A} \subseteq \mathcal{B} \models \mathsf{T} \implies \mathcal{A} \models \mathsf{T}.$$

Allgemeiner nennt man eine Formel $\varphi(x_1,\ldots,x_n)$ **Substruktur-persistent** in $\mathsf{T}$ (oder einfach $\mathsf{T}$-**persistent**) gdw

für alle $\mathcal{A},\mathcal{B} \models \mathsf{T}$ mit $\mathcal{A} \subseteq \mathcal{B}$ und für alle $a_1,\ldots,a_n \in A$ gilt:

$$\mathcal{B} \models \varphi[a_1,\ldots,a_n] \Longrightarrow \mathcal{A} \models \varphi[a_1,\ldots,a_n].$$

$\forall$-Formeln bleiben offensichtlich gültig, wenn man den Grundbereich einschränkt (und entsprechend bleiben existentielle Formeln wahr unter Erweiterungen), sie sind also persistent. Es gilt aber auch die Umkehrung:

### 3.4.3 Modellerweiterungssatz von Keisler

*Folgende Aussagen sind äquivalent für eine $\mathcal{L}$-Theorie* $\mathsf{T}$:

(i) Es existiert ein Struktur $\mathcal{B}$ mit $\mathcal{A} \subseteq \mathcal{B}$ und $\mathcal{B} \models \mathsf{T}$,

(ii) $\mathcal{A} \models \mathsf{T}_\forall$.

*Beweis* von (i) $\Rightarrow$ (ii):

Sei $\mathcal{B}$ mit $\mathcal{A} \subseteq \mathcal{B}$ und $\mathcal{B} \models \mathsf{T}$. Dann ist auch $\mathcal{B} \models \mathsf{T}_\forall$, und weil $\forall$-Sätze beim Übergang zu Substrukturen erhalten bleiben, ist $\mathcal{A} \models \mathsf{T}_\forall$.

(ii) $\Rightarrow$ (i): Sei $\mathcal{A} \models \mathsf{T}_\forall$. Wir behaupten, daß

$$\mathsf{T} \cup \mathsf{D}(\mathcal{A})$$

ein Modell hat: Anderenfalls gäbe es nach dem Kompaktheitssatz eine endliche Teilmenge $\Delta \subseteq \mathsf{D}(\mathcal{A})$, so daß $\mathsf{T} \cup \Delta$ kein Modell hat. Es sei $\delta(\underline{a_1},\ldots,\underline{a_n})$ die Konjunktion der endlich-vielen Sätze in $\Delta$, die dann in keinem Modell von $\mathsf{T}$ wahr sein können, d. h. es gilt

$$\mathsf{T} \models \neg\delta(\underline{a_1},\ldots,\underline{a_n}),$$

und da die Konstanten $\underline{a_i}$ in der Sprache von $\mathsf{T}$ nicht vorkommen, nach dem Satz über neue Konstanten 3.1.1 auch

$$\mathsf{T} \models \forall x_1,\ldots,x_n \neg\delta(x_1,\ldots,x_n),$$

also $\forall x_1,\ldots,x_n \neg\delta(x_1,\ldots,x_n) \in \mathsf{T}_\forall$. Da nach Voraussetzung $\mathcal{A} \models \mathsf{T}_\forall$, so:

$$\mathcal{A} \models \forall x_1,\ldots,x_n \neg\delta(x_1,\ldots,x_n), \quad \text{insbesondere}$$
$$\mathcal{A} \models \neg\delta[a_1,\ldots,a_n]$$

im Widerspruch zu $\Delta \subseteq \mathsf{D}(\mathcal{A})$! Somit hat also $\mathsf{D}(\mathcal{A}) \cup \mathsf{T}$ ein Modell, welches nach dem Diagrammlemma 3.4.1 als Oberstruktur von $\mathcal{A}$ gewählt werden kann. □

Als Korollare erhalten wir:

### 3.4.4 Erhaltungssatz für universelle Formeln

*Die folgenden Aussagen sind für einen Formel $\varphi$ äquivalent:*

(i) $\varphi = \varphi(x_1, \ldots, x_n)$ *ist* $\mathsf{T}$*-persistent,*

(ii) *es gibt eine universelle Formel* $\psi(x_1, \ldots, x_n)$*, die* $\mathsf{T}$*-äquivalent zu* $\varphi$ *ist, d. h.*

$$\mathsf{T} \models \varphi \leftrightarrow \psi.$$

*Beweis* von (i) $\Rightarrow$ (ii): Wir nehmen zunächst an, daß $\varphi$ ein Satz $\sigma$ ist, welcher $\mathsf{T}$-persistent ist. Setze

$$\Sigma := \{\varphi \mid \varphi \ universell, \mathsf{T} + \sigma \models \varphi\} = (\mathsf{T} + \sigma)_\forall.$$

Dann gilt

$$(*) \quad \mathsf{T} \cup \Sigma \models \sigma,$$

denn ist $\mathcal{A} \models \mathsf{T} \cup \Sigma$, so existiert nach dem Modellerweiterungssatz 3.4.3 ein $\mathcal{B}$ mit $\mathcal{A} \subseteq \mathcal{B} \models \mathsf{T} + \sigma$, also auch $\mathcal{A} \models \sigma$ nach (i).

Wegen (*) gibt es nach dem Kompaktheitssatz endlich-viele Sätze $\sigma_1, \ldots, \sigma_k \in \Sigma$ mit

$$\mathsf{T} + \sigma_1 \wedge \ldots \wedge \sigma_k \models \sigma,$$

und da auch umgekehrt $\mathsf{T} + \sigma \models \sigma_i$ für alle $i = 1, \ldots, k$, so gilt

$$\mathsf{T} \models \sigma \leftrightarrow \sigma_1 \wedge \ldots \wedge \sigma_k.$$

Die Konjunktion endlich-vieler universeller Sätze ist aber äquivalent zu einem universellen Satz (s. im nächsten Kapitel den Abschnitt 4.5). Den allgemeinen Fall einer Formel kann man auf den obigen Fall eines Satzes durch die Einführung neuer Konstanten und den Übergang zum Satz $\varphi(c_1, \ldots, c_n)$ zurückführen. $\square$

### 3.4.5 Satz von Łoś-Tarski

$$\begin{aligned} \mathsf{T} \text{ Substruktur-abgeschlossen} &\iff Mod(\mathsf{T}) = Mod(\mathsf{T}_\forall) \\ &\iff \mathsf{T} \text{ hat ein Axiomensystem aus } \mathsf{T}_\forall\text{-Sätzen.} \end{aligned}$$

## 3.5 Einige mathematische Theorien

### 3.5.1 Ordnungen

Die Theorie LO der *partiellen Ordnungen (Halbordnungen)* wird in einer Sprache mit einem 2-stelligen Relationszeichen $<$ formuliert und hat folgende Axiome:

| | | |
|---|---|---|
| (L1) | $\forall x\, x \not< x$ | **irreflexiv** |
| (L2) | $\forall x \forall y \forall z\, (x < y \wedge y < z \rightarrow x < z)$ | **transitiv** |

Jede Menge $A$ läßt sich partiell ordnen durch $\{x, y \mid x, y \in A \wedge x \neq y\}$, wobei alle Elemente miteinander unvergleichbar sind. Dagegen erfüllen *linearen Ordnungen* zusätzlich das Axiom

| | | |
|---|---|---|
| (L3) | $\forall x \forall y\, (x < y \vee x = y \vee y < x)$ | **connex** |

so daß alle Elemente miteinander vergleichbar sind. Da sich jede partielle Ordnung auf einer endlichen Menge zu einer linearen Ordnung erweitern läßt, folgt mit Hilfe des Kompaktheitssatzes:

- *Jede partielle Ordnung auf einer Menge läßt sich zu einer linearen Ordnung erweitern,* insbesondere:
- *jede Menge besitzt eine lineare Ordnung.*

Dieses Ergebnis ist keineswegs trivial, so kann man z. B. für die Menge der reellen Funktionen zwar eine partielle Ordnung leicht angeben, nicht aber eine lineare Ordnung (zumindest nicht ohne besondere Zusatzaxiome). Der Kompaktheitssatz (und damit die Existenz linearer Ordnungen) folgt aus dem mengentheoretischen *Auswahlaxiom* (siehe 7.2), welches sogar die Existenz von Wohlordnungen auf beliebigen Mengen nach sich zieht: Eine *Wohlordnung* auf einer Menge $A$ ist eine lineare Ordnung, in welcher jede nicht-leere Teilmenge ein kleinstes Element besitzt (siehe 6.1.1). Diese Forderung ist nun nicht mehr als Eigenschaft 1. Stufe (wobei also nur über die Elemente einer Menge quantifiziert wird) ausdrückbar, was sich leicht mit Hilfe des Kompaktheitssatzes zeigen läßt. Dagegen lassen sich *dichte* Ordnungen axiomatisieren; die entsprechenden Theorien werden wir gesondert in 9.2 behandeln.

### 3.5.2 Gruppen- und Körpertheorie

1. Die Theorie der (ABELschen) Gruppen in der Sprache $\mathcal{L}_G$ mit einem 2-stelligen Funktionssymbol $\circ$, einem 1-stelligen Funktionssymbol $^{-1}$ und der Individuenkonstanten $e$ besitzt universelle Axiome und ist somit Substruktur-abgeschlossen. Ist dagegen $\mathsf{T}_{\mathsf{AG}}$ die Theorie der ABELschen Gruppen in der Sprache mit nur einem 2-stelligen Funktionssymbol $\circ$, dann ist $\mathsf{T}_\forall$ die Theorie der kommutativen regulären Halbgruppen mit den Axiomen:

- $\forall x \forall y \forall z\, x \circ (y \circ z) = (x \circ y) \circ z$ (assoziativ)
- $\forall x \forall y\, x \circ y = y \circ x$ (kommutativ)
- $\forall x \forall y \forall z (x \circ y = x \circ z \rightarrow y = z)$ (Kürzungsregel)

In ähnlicher Weise ist für die Theorie $\mathsf{T}_\mathsf{F}$ der Körper (in der Sprache mit $+, \cdot, -, 0, 1$) der universelle Teil $(\mathsf{T}_\mathsf{F})_\forall$ die Theorie der Integritätsbereiche.

2. Die Klasse der *unendlichen* Gruppen, Körper, ... läßt sich mit Hilfe der (unendlichen) Satzmenge $\Sigma_\infty$ axiomatisieren, nicht aber die Klasse der *endlichen* Gruppen, der *endlichen* Körper, ... (denn es gibt jeweils beliebig große endliche Gruppen und Körper, und eine Theorie dieser Modelle muß dann nach dem Satz 3.3.2 über die Existenz unendlicher Modelle notwendig ein unendliches Modell einschließen). Ähnliches gilt für die Theorie der Körper der Charakteristik 0 bzw. endlicher Charakteristik:

#### Satz

*$\mathcal{L}$ sei die Sprache der Theorie der Körper, $\sigma$ ein Satz dieser Sprache.*
*Gilt $\sigma$ in allen Körpern der Charakteristik 0 (bzw. in allen unendlichen Körpern), so gibt es ein p, so daß $\sigma$ in allen Körpern der Charakteristik $\geq p$ (bzw. in allen Körpern mit mindestens p Elementen) gilt.*

*Beweis:* Es sei $\mathsf{T}_\mathsf{F}$ die Theorie der Körper ($\mathsf{F}$ für englisch: *field*). Nimmt man den Satz

$$\chi_p := \underbrace{1 + 1 + \ldots + 1}_{p-mal} = 0$$

hinzu, welcher ausdrückt, daß die Charakteristik = $p$ ist, so erhält man die Theorie der Körper der Charakteristik $p$. Um die Theorie $\mathsf{T}_{\mathsf{F},0}$ der Körper der Charakteristik 0 zu erhalten, muß man $\mathsf{T}_\mathsf{F}$ ergänzen um die unendlich-vielen Aussagen

$$\{\neg\chi_p \mid p \; Primzahl\}.$$

Gilt nun $\sigma$ in allen Körpern der Charakteristik 0, so $\mathsf{T}_{\mathsf{F},0} \models \sigma$, also gibt es nach dem Kompaktheitssatz ein endliches $\mathsf{T}_0 \subseteq \mathsf{T}_{\mathsf{F},0}$ mit $\mathsf{T}_0 \models \sigma$. $\mathsf{T}_0$ kann aber nur endlich viele der Aussagen $\neg\chi_p$ enthalten, und daraus erhalten wir die Behauptung. (Ähnlich im anderen Fall mit Anzahl- statt Charakteristik-Formeln.) $\square$

Mit modelltheoretischen Methoden lassen sich noch weitere Ergebnisse über verschiedene algebraische Theorien erzielen, insbesondere die *Entscheidbarkeit* der Theorie von $(\mathbb{R},<,+,-,\cdot,0,1)$ (= Theorie der reell-abgeschlossenen Körper) und der Theorie von $(\mathbb{C},+,-,\cdot,0,1)$ (= Theorie der algebraisch-abgeschlossenen Körper).

### 3.5.3 Axiomatisierbarkeit: Elementare Klassen

Ist $K$ eine Klasse von $\mathcal{L}$-Strukturen, so heißt

$$\begin{array}{rcl} K \textbf{ elementar} \text{ (EC-Klasse)}: & \iff & K = Mod(\sigma) \text{ für einen Satz } \sigma, \\ K\, \Delta\textbf{-elementar} \text{ (EC}_\Delta\text{-Klasse)}: & \iff & K = Mod(\Gamma) \text{ für eine Satzmenge } \Gamma. \end{array}$$

Eine EC-Klasse ist also eine Klasse von Strukturen, die sich durch einen einzelnen Satz (oder durch endlich-viele Sätze) axiomatisieren läßt, eine $\mathrm{EC}_\Delta$-Klasse ist Durchschnitt von EC-Klassen und läßt sich mittels einer Menge von Sätzen axiomatisieren.

**Beispiele**

1. elementare Klassen sind:
   - $\{\mathcal{G} \mid \mathcal{G} \textit{ ist Gruppe}\} = Mod(\sigma_G)$,
     $\sigma_G$ = Konjunktion der endlich-vielen Gruppenaxiome,
   - $\{\mathcal{G} \mid \mathcal{G} \textit{ ist kommutative Gruppe}\}$,
   - $\{\mathcal{R} \mid \mathcal{R} \textit{ ist Ring (kommutativ, mit 1)}\}$,
   - $\{\mathcal{K} \mid \mathcal{K} \textit{ ist Körper}\}$,
   - $\{\mathcal{K} \mid \mathcal{K} \textit{ ist Körper der Charakteristik } p\}$ ($p$ fest; $p$ Primzahl),
   - $\{\leq \mid \leq \textit{ ist (partielle, lineare) Ordnung}\}$.
2. $\Delta$-elementare Klassen, aber *keine elementaren Klassen* sind:
   - $\{\mathcal{G} \mid \mathcal{G} \textit{ ist unendliche Gruppe}\} = Mod(\{\sigma_G\} \cup \Sigma_\infty)$,

- $\{\mathcal{K} \mid \mathcal{K}$ *ist unendlicher Körper*$\}$,
- $\{\mathcal{K} \mid \mathcal{K}$ *ist Körper der Charakteristik 0*$\}$.

3. *nicht* einmal $\Delta$-elementare Klassen sind:

- $\{\mathcal{G} \mid \mathcal{G}$ *ist endliche Gruppe*$\}$,
- $\{\mathcal{G} \mid \mathcal{G}$ *ist Torsions-Gruppe (alle Elemente haben endliche Ordnung)*$\}$,
- $\{\mathcal{K} \mid \mathcal{K}$ *ist endlicher Körper*$\}$,
- $\{\leq \mid \leq$ *ist Wohlordnung*$\}$.

Die negativen Beispiele beruhen auf dem Kompaktheitssatz und dieser wiederum wesentlich auf der Eigenschaft von Beweisen, endlich zu sein. Geht man zu erweiterten Logiken über, die etwa auch unendliche Konjunktionen und Disjunktionen zulassen, so kann man zwar wesentlich mehr ausdrücken, aber Beweise sind nun nicht mehr notwendig endlich, die erweiterte Logik ist nicht mehr vollständig (bzw. nur in einem abgeschwächten Sinn), auch gelten die klassische Sätze der Modelltheorie dann höchstens noch mit Einschränkungen.

### 3.5.4 Zahlentheorie

Während die algebraischen Theorien der Gruppen, Ringe, Körper, ... sehr viele Modelle besitzen, untersucht die Zahlentheorie Eigenschaften *eines* Modells, das man das **Standardmodell** der Zahlentheorie nennt:

$$\mathcal{N} = (\mathbb{N}, +, \cdot, ', 0).$$

$\mathcal{L}_{\mathsf{PA}}$ sei die zugehörige Sprache, für die wir auch $+, \cdot, 0$ als (nicht-logische) Symbole wählen, aber $S$ für die 1-stellige Nachfolgeroperation. Die Theorie des Standardmodells, $\mathsf{Th}(\mathcal{N})$, ist die Menge der $\mathcal{L}_{\mathsf{PA}}$-Sätze, die in diesem Modell gelten - was sich zwar einfach definieren läßt, aber doch immer noch viele offene Fragen bereithält. Die wichtigsten Eigenschaften des Standardmodells werden zusammengefaßt als

**Peano-Axiome:**

P1 $Sx \neq 0$

P2 $Sx = Sy \rightarrow x = y$

P3 $x + 0 = x$

P4 $x + Sy = S(x + y)$

P5 $x \cdot 0 = 0$

P6 $x \cdot Sy = x \cdot y + x$

sowie die unendlich-vielen **Induktionsaxiome**:

Ind $\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(Sx)) \to \forall x\, \varphi(x)$

Die Theorie mit diesen Axiomen nennt man PA, die **Peano-Arithmetik**. Aus dem Kompaktheitssatz folgt nun, daß diese Theorie **Nicht-Standardmodelle** besitzt, d. h. Modelle, die nicht isomorph zum Standardmodell sind, weil sie "unendlich-große" Zahlen besitzen:

**Satz**

*Es gibt ein (abzählbares) Modell* $\mathcal{N}^*$ *von* PA*, welches nicht-isomorph zu* $\mathcal{N}$ *ist.*

(Tatsächlich gibt es - wie wir später zeigen werden - sogar in jeder unendlichen Kardinalzahl ein Nicht-Standardmodell.)

*Beweis:* Erweitere die Sprache von PA um eine neue Konstante $c$ für ein "unendlich-großes" Element und bilde die erweiterte Theorie

$$\mathsf{PA}^* := \mathsf{PA} \cup \{c \neq 0, c \neq S0, c \neq SS0, \ldots\}.$$

Da es unendlich-viele natürliche Zahlen gibt, kann man endlich-viele der zusätzlichen Aussagen im Standardmodell erfüllen, indem man $c$ als genügend große Zahl wählt. Somit hat jede endliche Teiltheorie von $\mathsf{PA}^*$ ein Modell, also gibt es nach dem Kompaktheitssatz auch ein Modell $\mathcal{N}^* \models \mathsf{PA}^*$. Da $\mathcal{N}^*$ die Axiome (P1) und (P2) erfüllt, ist die Abbildung der Terme $\{0, S0, SS0, \ldots\}$ auf ihre Interpretationen in $\mathcal{N}^*$ injektiv, so daß wir sie mit den "gewöhnlichen" natürlichen Zahlen identifizieren können. Dann hat aber $\mathcal{N}^*$ noch das Element $c^{\mathcal{N}^*}$, welches von allen "gewöhnlichen" natürlichen Zahlen verschieden sein muß (da nämlich $\mathcal{N}^*$ auch die neuen Axiome von $\mathsf{PA}^*$ über $c$ erfüllt). $\square$

Dasselbe Ergebnis erhalten wir auch, wenn wir die Theorie PA durch die Theorie des Standardmodells ersetzen; wieder gibt es ein nicht-isomorphes (abzählbares) Modell $\mathcal{N}^*$, welches aber alle Sätze erfüllt, die auch im Standardmodell gelten. Ein solches Modell kann man nach A. ROBINSON benutzen, um eine Nonstandard-Analyis aufzubauen, d. h. eine Theorie, in welcher alle Aussagen der “gewöhnlichen” reellen Zahlen gelten, in denen es aber noch zusätzlich “infinitesimal kleine” Zahlen gibt.

Mittels Aussagen in der Sprache $\mathcal{L}_{\mathsf{PA}}$ kann man daher das Standardmodell nicht charakterisieren, offenbar ist also die Sprache der Prädikatenlogik nicht ausdrucksstark genug. Es liegt daher nahe, das Induktionsschema zu verstärken, indem man den Eigenschaftsbegriff erweitert, und zwar geht man über zu einer *Sprache 2. Stufe*, in welcher man zusätzlich die Möglichkeit hat, mittels *Quantoren 2. Stufe* $\exists P, \forall P$ über alle *Teilmengen* (statt nur über die Elemente) des jeweiligen Grundbereiches zu quantifizieren:

### Peano-Axiome 2. Stufe

Die Theorie $\mathsf{PA}^{(2)}$ (PEANO-Axiome 2. Stufe) besitzt folgende Axiome:

P1 $\quad \forall x\, Sx \neq 0$

P2 $\quad \forall x \forall y (Sx = Sy \rightarrow x = y)$

IND $\quad \forall P[P(0) \wedge \forall x(P(x) \rightarrow P(Sx)) \rightarrow \forall x\, P(x)]$

Allerdings taucht nun ein neues Problem auf: Im Gegensatz zur gewöhnlichen Logik der ersten Stufe gibt es für die Logik 2. Stufe kein vollständiges Axiomensystem (zumindest solange nicht, wie man verlangt, daß der Begriff eines Axioms entscheidbar ist) und auch die meisten anderen Ergebnisse lassen sich nicht auf die erweiterte Sprache übertragen (tatsächlich ist es ja gerade der Kompaktheitssatz, der zu Nichtstandardmodellen führt). Nun gibt es aber einen Ausweg: Wir können die obigen Axiome mengentheoretisch deuten und werden zeigen (6.3.1), daß im Rahmen der Mengenlehre das obige Axiomensystem genau ein Modell hat bis auf Isomorphie, die Theorie also kategorisch ist in folgendem Sinn:

*Jedes Modell von* $\mathsf{PA}^{(2)}$ *ist isomorph zum Standardmodell* $(\mathbb{N},', 0)$.

Die Mengenlehre ihrerseits werden wir in der Sprache der Logik erster Stufe axiomatisieren, so daß diese Theorie ihrerseits nicht kategorisch ist, also nicht-isomorphe Modelle besitzt.

### 3.5.5 Mengenlehre

Die Mengenlehre spielt (ähnlich wie die Zahlentheorie) in der Mathematischen Logik eine besondere Rolle:

1. Als *Grundlagentheorie* stellt sie sämtliche Objekte, die in den einzelnen mathematischen Disziplinen untersucht werden:

   Zahlen, Funktionen, Operatoren, Relationen, Punkte, Räume ...

   unter dem einheitlichen Begriff der *Menge* dar, insbesondere haben wir mengentheoretische Begriffsbildungen bei der Einführung des Strukturbegriffes (inhaltlich) bereits benutzt. Während syntaktische Begriffe auf zahlentheoretischer Grundlage eingeführt werden können[2], benötigt die Semantik der Prädikatenlogik bereits den allgemeinen Mengenbegriff; eine geeignete Fassung der Mengenlehre gehört somit zur *Metamathematik* der Prädikatenlogik.

2. Außerdem ist die Mengenlehre selbst eine formale *mathematische Theorie* des Unendlichen (insbesondere der transfiniten Ordinal- und Kardinalzahlen). (Eine ausführlichere Darstellung folgt im Teil II.)

Beide Aspekte lassen sich nicht völlig voneinander trennen; insbesondere das Auswahlaxiom (mit seinen vielen äquivalenten Fassungen) hat zahlreiche Anwendungen in fast allen mathematischen Gebieten (einschließlich der Mathematischen Logik), ist aber auch wichtig für die Entwicklung der Theorie der transfiniten Kardinalzahlen.

[2] Darauf werden wir bei der Behandlung der GÖDELschen Unvollständigkeitssätze im Teil IV genauer eingehen.

# Kapitel 4

# Gesetze der Prädikatenlogik

## 4.1 Ein Axiomensystem für die Prädikatenlogik

Wir haben gesehen, daß es für den aussagenlogischen Wahrheitsbegriff, ag[$\varphi$], verschiedene Entscheidungsverfahren gibt und für endliche Formelmengen $\mathsf{T}$ der aussagenlogische Folgerungsbegriff, $\mathsf{T} \models \varphi$, entscheidbar ist, da man nur *endlich-viele* Belegungen der Aussagenvariablen nachzuprüfen hat. Dagegen erfordert der Wahrheitsbegriff $\models \varphi$ (und ebenso der semantische Folgerungsbegriff $\mathsf{T} \models \varphi$) der Prädikatenlogik den Nachweis der Gültigkeit von $\varphi$ in allen Strukturen der zugehörigen Sprache bzw. in allen Modellen von $\mathsf{T}$, und hierzu hat man nicht nur sehr viele Strukturen, sondern unter diesen i. a. auch noch *unendliche* Strukturen zu untersuchen, so daß wir in diesem Falle keine Entscheidbarkeit erwarten können. Daher ist es von Bedeutung, daß es auch einen syntaktischen Beweisbegriff, $\mathsf{T} \vdash \varphi$, gibt, welcher mit dem semantischen übereinstimmt (verallgemeinerter Vollständigkeits- und Korrektheitssatz). Damit erhalten wir zwar kein Entscheidungsverfahren, aber immerhin die Möglichkeit, alle Folgerungen aus einer Theorie in standardisierter Weise aufzuzählen. Außerdem zeigt ein vollständiges Axiomensystem mit seinen Axiomen und Regeln an, auf welche grundlegende Prinzipien der Folgerungsbegriff zurückzuführen ist.

Beim Übergang von der Aussagenlogik zur Prädikatenlogik kommen zu den aussagenlogischen Symbolen hinzu:

- atomare Formeln (anstelle der Aussagenvariablen),
- die Quantoren (wobei wir den $\forall$-Quantor mittels $\exists$ definieren) sowie
- die Gleichheit = (als 2-stelliges logisches Relationszeichen).

Entsprechend werden die aussagenlogischen Axiome und Regeln ergänzt um zusätzliche Axiome und Regeln für die neuen logischen Symbole.

### 4.1.1 Axiomensystem von Shoenfield

Wie im Falle der Aussagenlogik gibt es verschiedene vollständige und korrekte Axiomensysteme für die Prädikatenlogik; wir wollen hier wieder ein spezielles Axiomensystem nach SHOENFIELD behandeln, dessen aussagenlogischen Teil wir bereits in 1.5.3 (b) angegeben haben:

**Axiome**:

(1) aussagenlogische Axiome: $\neg\varphi \vee \varphi$

(2) Substitutionsaxiome: $\varphi[t/x] \to \exists x\varphi$, falls $t$ in $\varphi$ substituierbar ist.

(3) Gleichheitsaxiome:

1. $x = x$
2. $x_1 = y_1 \wedge \ldots \wedge x_m = y_m \to Fx_1 \ldots x_m = Fy_1 \ldots y_m$
   für jedes $m$-stellige Funktionszeichen $F$
3. $x_1 = y_1 \wedge \ldots \wedge x_n = y_n \ \wedge \ Rx_1 \ldots x_n \to \ Ry_1 \ldots y_n$
   für jedes $n$-stellige Relationszeichen $R$

Das letzte Gleichheitsaxiom wird insbesondere auch für die 2-stellige Gleichheitsrelation gefordert!

Das Substitutionsaxiom besagt, daß man eine Existenzaussage nachweisen kann, indem man ein *Beispiel* angibt.

**Regeln:**

Zu den uns bereits bekannten aussagenlogischen Regeln

Expansion: $\dfrac{\psi}{\varphi \vee \psi}$ Assoziativität: $\dfrac{\varphi \vee (\psi \vee \delta)}{(\varphi \vee \psi) \vee \delta}$

Kürzung: $\dfrac{\varphi \vee \varphi}{\varphi}$ Schnitt: $\dfrac{\varphi \vee \psi, \neg\varphi \vee \delta}{\psi \vee \delta}$

kommt eine Quantorenregel hinzu:

$\exists$-Einführung: $\dfrac{\varphi \to \psi}{\exists x\varphi \to \psi}$, falls x in $\psi$ nicht frei vorkommt.

**Bemerkungen**

1. Die Regel der $\exists$-Einführung heißt auch **Vordere Partikularisierung**. Die angefügte **Variablenbedingung** ist notwendig, sonst würde man aus der beweisbaren Formel $x = y \to x = y$ durch $\exists$-Einführung die Formel $\exists x(x = y) \to x = y$ und mit der beweisbaren Formel $\exists x(x = y)$ auch die Formel $x = y$ als beweisbar erhalten, welche aber nicht wahr ist, da sie nur in 1-elementigen Strukturen gilt. (Inhaltlich kann man auch so argumentieren: Es gilt $x = 3 \to x$ ist Primzahl, aber nicht $\exists x(x = 3) \to x$ ist Primzahl!)

2. Wo benutzt man die $\exists$-Einführung? Nehmen wir an, daß wir in einem Beweis zu der Aussage $\varphi(x) \to \psi$ gelangt sind. Um hieraus die Aussage $\psi$ zu erhalten, brauchen wir - wenn $\psi$ von der Wahl eines $x$ nicht abhängt (formal: wenn $x$ in $\psi$ nicht frei vorkommt) - nur zu versuchen, die Voraussetzung $\varphi(x)$ für (mindestens) ein $x$ nachzuweisen (formal: die Aussage $\exists x\varphi(x)$ zu beweisen). Dann können wir nämlich mittels $\exists$-Einführung und modus ponens tatsächlich die gewünschte Aussage $\psi$ erhalten.

Den Beweisbegriff für ein Axiomensystem der Prädikatenlogik (wie das oben angegebene System von SHOENFIELD) können wir wörtlich vom Aussagenkalkül übernehmen:

### 4.1.2 Definition eines Beweises

Ein **Beweis** von $\varphi$ aus $\mathsf{T}$ ist eine endliche Folge $(\varphi_1, \dots, \varphi_n)$ von Formeln, so daß

(B1) $\varphi_n = \varphi$,

(B2) für jedes $m \leq n$ gilt:

(1) $\varphi_m$ ist Axiom *oder*

(2) $\varphi_m$ ist Formel aus $\mathsf{T}$ *oder*

(3) es gibt ein $i < m$, so daß $\varphi_m$ aus $\varphi_i$ hervorgeht durch Anwendung einer Regel *oder*

(4) es gibt $i, j < m$, so daß $\varphi_m$ aus $\varphi_i$ und $\varphi_j$ hervorgeht durch Anwendung einer Regel mit 2 Prämissen.

Wir schreiben wieder:

| | | |
|---|---|---|
| $\mathsf{T} \vdash \varphi :\Longleftrightarrow$ | es gibt einen Beweis von $\varphi$ aus $\mathsf{T}$ | $\varphi$ ist aus $\mathsf{T}$ **beweisbar**, |
| $\vdash \varphi :\Longleftrightarrow$ | $\emptyset \vdash \varphi$ | $\varphi$ ist **beweisbar**. |

**Bemerkung**

Die Beweisbarkeit läßt sich auch rekursiv wie folgt charakterisieren:

(S1) Jedes logische Axiom und jedes Axiom von $\mathsf{T}$ sind aus $\mathsf{T}$ beweisbar.

(S2) Ist die Prämisse (bzw. sind die Prämissen) einer Regel aus $\mathsf{T}$ beweisbar, so auch ihre Konklusion.

**Satz über die Verallgemeinerte Korrektheit**

$$\mathsf{T} \vdash \varphi \Longrightarrow \mathsf{T} \models \varphi$$

*Beweisskizze:* Es sei $\mathcal{A} \models \mathsf{T}$. Zeige

$$(*) \quad \mathsf{T} \vdash \varphi \Longrightarrow \mathcal{A} \models \varphi$$

durch Induktion gemäß der induktiven Definition (S1), (S2):

1. $\varphi$ ist logisches Axiom. Zeige, daß dann $\mathcal{A} \models \varphi$.
2. $\varphi$ ist Axiom von $\mathsf{T}$. Dann gilt die Beh. (*) offenbar wegen $\mathcal{A} \models \mathsf{T}$.
3. $\varphi$ ist Konklusion einer Regel, deren Prämisse(n) aus $\mathsf{T}$ beweisbar ist (sind).

Gehe die einzelnen Regeln durch und zeige, daß sich Gültigkeit in $\mathcal{A}$ von den Prämissen auf die Konklusion vererbt. □

## 4.2 Der Tautologiesatz

Hier wollen wir erläutern, in welcher Weise das Axiomensystem von SHOENFIELD die Aussagenlogik enthält und "aussagenlogisch vollständig" ist (Tautologiesatz):

Den Grundbestandteilen der aussagenlogischen Sprache, also den Aussagenvariablen, entsprechen in der prädikatenlogischen Sprache die atomaren Formeln (Primformeln), während die aussagenlogischen Operationen hier selbst wieder vorkommen. Andererseits gibt es prädikatenlogische Formeln, die nicht atomar, aber auch aussagenlogisch nicht weiter zerlegbar sind; diese werden zusammen mit den atomaren Formeln *elementar* genannt:

**Definition einer Tautologie**

$$\begin{aligned} \varphi \quad \textbf{elementar} \quad &:\Longleftrightarrow \ \varphi \text{ atomar oder beginnt mit dem Quantor } \exists. \\ B \ \textbf{Bewertung} \text{ von } \mathcal{L} &:\Longleftrightarrow \ B : \{\varphi \mid \varphi \, \text{elementare } \mathcal{L}\text{-Formel}\} \longrightarrow \{W, F\}. \end{aligned}$$

Die Bewertungen der elementaren Formeln treten an die Stelle der Belegungen $f$ der Aussagenvariable in der Aussagenlogik; ebenso wie sich diese (mittels der Wahrheitstafeln) fortsetzen lassen zum Wahrheitswert $f(\varphi)$ einer aussagenlogischen Formel $\varphi$, können wir jede Bewertung $B$ fortsetzen zu einer Abbildung

$$B : \{\varphi \mid \varphi \ \mathcal{L}\text{-Formel}\} \longrightarrow \{W, F\}$$

durch die Festlegungen

$$\begin{aligned} B(\neg\varphi) = W \quad &\Longleftrightarrow \quad B(\varphi) = F, \\ B(\varphi \vee \psi) = W \quad &\Longleftrightarrow \quad B(\varphi) = W \ \mathit{oder} \ B(\psi) = W. \end{aligned}$$

An die Stelle der entsprechenden semantischen Begriffe der Aussagenlogik treten nun die folgenden:

$$\begin{aligned} \models_{AL} \varphi \!: &\Longleftrightarrow \ (B(\varphi) = W \text{ für alle Bewertungen } B) \qquad \varphi \text{ ist } \textbf{Tautologie}, \\ \mathsf{T} \models_{AL} \varphi \!: &\Longleftrightarrow \ (B(\varphi) = W \text{ für alle Bew. } B \text{, für die } B(\psi) = W \text{ für alle } \psi \in \mathsf{T}) \\ & \qquad\qquad \varphi \text{ ist } \textbf{tautologische Folgerung} \text{ aus } \mathsf{T}. \end{aligned}$$

**Beispiele**

Tautologien sind: $\varphi \vee \neg\varphi$, $\varphi \vee \psi \leftrightarrow \psi \vee \varphi$, nicht aber (obwohl wahre Aussagen): $x = x$, $x = y \rightarrow y = x$, $\exists x(x = x)$; tatsächlich sind die letzten Formeln nicht aussagenlogisch, sondern erst prädikatenlogisch wahr (als elementare Formeln könnten sie einfach mit $F$ belegt werden).

**Tautologiesatz**

(i) *Falls* $\mathsf{T} \vdash \varphi_1, \ldots, \varphi_n$ *und* $\varphi_1, \ldots, \varphi_n \models_{AL} \varphi$, *so* $\mathsf{T} \vdash \varphi$,

*(ii) falls* $\mathsf{T}$ *endlich und* $\mathsf{T} \models_{AL} \varphi$, *so* $\mathsf{T} \vdash \varphi$,

*(iii) falls* $\models_{AL} \varphi$, *so* $\vdash \varphi$, *und somit* $\mathsf{T} \vdash \varphi$ *für jede Tautologie* $\varphi$.

*Beweis:* Offensichtlich folgt (i) aus (ii) , und (ii) folgt aus (iii): Sei

| | |
|---|---|
| $\mathsf{T} = \{\varphi_1, \ldots, \varphi_n\}$ und $\mathsf{T} \models_{AL} \varphi$, | also |
| $\varphi_1, \ldots, \varphi_n \models_{AL} \varphi$. | Dann gilt auch |
| $\models_{AL} \varphi_1 \to \ldots \to \varphi_n \to \varphi$, | also nach (iii) |
| $\vdash \varphi_1 \to \ldots \to \varphi_n \to \varphi$ | und daraus mit modus ponens auch |
| $\varphi_1, \ldots, \varphi_n \vdash \varphi$. | |

Somit bleibt nur (iii), d.h. die einfache Vollständigkeit des aussagenlogischen Teils, zu beweisen. Dazu geht man wie beim Beweis des Vollständigkeitssatzes für die Aussagenlogik vor und erhält

$$\models_{AL} \varphi \Longrightarrow \models A_\varphi,$$

wobei $A_\varphi$ die aussagenlogische Formel ist, die aus $\varphi$ entsteht, indem man die elementaren Formeln $\varphi_1, \ldots, \varphi_n$ durch Aussagenvariable $A_1, \ldots, A_n$ ersetzt. Der Vollständigkeitssatz für die Aussagenlogik ergibt nun $\vdash A_\varphi$, und aus einem aussagenlogischen Beweis von $A_\varphi$ erhält man einen (prädikatenlogischen) Beweis von $\varphi$, indem man in diesem Beweis nun umgekehrt die Aussagenvariable $A_1, \ldots, A_n$ durch die elementaren Formeln $\varphi_1, \ldots, \varphi_n$ ersetzt. □

Aufgrund der Aussage (i) des Tautologiesatzes können wir somit aussagenlogische Argumente (präziser: tautologische Folgerungen) in prädikatenlogische Folgerungen übernehmen, wovon wir im folgenden häufig stillschweigend Gebrauch machen werden.

## 4.3 Substitution und universeller Abschluß

Den $\forall$-Quantor haben wir definiert durch

$$\begin{array}{rcll} \forall x\varphi & \leftrightarrow & \neg\exists x\neg\varphi, & \text{somit gilt auch} \\ \exists x\varphi & \leftrightarrow & \neg\forall x\neg\varphi & \text{und} \\ \neg\forall x\varphi & \leftrightarrow & \exists x\neg\varphi & \text{und} \\ \neg\exists x\varphi & \leftrightarrow & \forall x\neg\varphi. & \end{array}$$

Die folgenden Sätze bezeichnen ableitbare Regeln bzw. beweisbare Formeln, die somit bei Beweisen verwendet werden können:

### 4.3.1 Hintere Generalisierung

$$\varphi \to \psi \vdash \varphi \to \forall x \psi,$$

*falls x in $\varphi$ nicht frei vorkommt (Variablen-Bedingung).*

*Beweis:*

| | | |
|---|---|---|
| (1) | $\varphi \to \psi$ | Voraussetzung |
| (2) | $\neg\psi \to \neg\varphi$ | tautologisch aus (1) |
| (3) | $\exists x \neg\psi \to \neg\varphi$ | mit $\exists$-Einführung (die Var.bedingung ist erfüllt!) |
| (4) | $\varphi \to \neg\exists x \neg\psi$ | tautologisch aus (3) |
| | $\varphi \to \forall x \psi$ | nach Definition des $\forall$-Quantors □ |

Diese Regel läßt sich in folgendem Fall anwenden:

> Zu beweisen sei die Aussage $\varphi \to \forall x \psi$.
> Dazu setzen wir $\varphi$ voraus und zeigen die Behauptung $\forall x \psi$ wie folgt: Sei "$x$ beliebig" (das heißt inhaltlich: wir dürfen kein $x$ wählen, für das $\varphi$ bereits eine besondere Eigenschaft festlegt - wie z. B. "$x$ ist Primzahl" oder "$x$ ist das größte Element", formal haben wir die Bedingung, daß $x$ nicht frei in $\varphi$ vorkommt). Können wir damit $\psi(x)$ beweisen, so haben wir nach der obigen Regel die Aussage $\varphi \to \forall x \psi$ gezeigt.

### 4.3.2 Satz über die Generalisierung

$$\varphi \vdash \forall x \varphi$$

*Beweis* :

| | | |
|---|---|---|
| (1) | $\varphi$ | Voraussetzung |
| (2) | $\neg\forall x\varphi \to \varphi$ | tautologisch: Abschwächung von (1) |
| (3) | $\neg\forall x\varphi \to \forall x\varphi$ | mit $\forall$-Einführung (die Var.bedingumg ist erfüllt!) |
| (4) | $\forall x\varphi$ | tautologisch aus (3) □ |

Das bedeutet also:

> Um $\forall x \varphi(x)$ aus T zu beweisen, genügt es $\varphi(x)$ (für ein "beliebiges $x$") aus T herzuleiten.

### 4.3.3 Substitutionsregel

$$\varphi(x_1,\dots,x_n) \vdash \varphi(t_1,\dots,t_n)$$

*Beweis* durch Induktion nach $n$:
$n = 1$:

| | | |
|---|---|---|
| | $\varphi$ | Voraussetzung |
| (1) | $\forall x_1 \varphi$ d.h. $\neg\exists x_1 \neg\varphi$ | Generalisierung |
| (2) | $\neg\varphi[t_1/x_1] \to \exists x_1 \neg\varphi$ | Substitutionsaxiom |
| (3) | $\neg\exists x_1 \neg\varphi \to \varphi[t_1/x_1]$ | tautologisch aus (2) |
| | $\varphi[t_1/x_1]$ | aus (1) und (3) mit modus ponens |

$n \geq 1$: Setze $\varphi' = \varphi[t_1/x_1,\dots,t_n/x_n]$ und wähle $n$ neue Variable $y_1,\dots,y_n$, die weder in $\varphi$ noch in $\varphi'$ vorkommen. (Dadurch erreichen wir, daß für die neuen Variablen die hintereinander ausgeführte Substitution mit der simultanen Substitution übereinstimmt.)

Wenden wir nun den Fall $n = 1$ sukzessive an, so erhalten wir aus $\varphi$:

$$\varphi[y_1/x_1],\ \varphi[y_1/x_1,y_2/x_2],\dots,\varphi[y_1/x_1,\dots,y_n/x_n]$$

und ebenso aus dem Fall $n = 1$:

$$\varphi[y_1/x_1,\dots,y_n/x_n][t_1/y_1] = \varphi[t_1/x_1,\dots,y_n/x_n],$$
$$\varphi[t_1/x_1,t_2/x_2,\dots,y_n/x_n],\dots,\varphi[t_1/x_1,t_2/x_2\dots,t_n/x_n]. \qquad \square$$

Die Bedeutung ist klar: Unter der Annahme $\varphi(x_1,\dots,x_n)$ gilt diese Aussage generell (Generalisierung), also insbesondere für beliebige (substituierbare) Terme (wobei diese wiederum neue Variable einführen können). - Als Verallgemeinerung des Substitutionsaxioms erhalten wir (zusammen mit seiner dualen Fassung (ii)):

### 4.3.4 Substitutionssatz

$$(i) \quad \varphi[t_1/x_1,\dots,t_n/x_n] \to \exists x_1 \dots \exists x_n \varphi$$
$$(ii) \quad \forall x_1 \dots \forall x_n \varphi \to \varphi[t_1/x_1,\dots,t_n/x_n]$$

(In beiden Fällen sei vorausgesetzt, daß $t_1,\dots,t_n$ für $x_1,\dots,x_n$ in $\varphi$ substituierbar sind.)

*Beweis* von (i): Aus den Substitutionsaxiomen folgt durch mehrfache Anwendung:

$$\varphi \to \exists x_n \varphi$$
$$\exists x_{i-1} \dots \exists x_n \to \exists x_i \exists x_{i-1} \dots \exists x_n \varphi$$
$$\varphi \to \exists x_1 \dots \exists x_n \varphi$$

und daraus folgt die Behauptung mit der Substitutionsregel.

(ii) beweist man analog (oder durch Kontraposition aus (i)). □

Beachte, daß die Umkehrungen von (i) und (ii), $\varphi(t_1,\dots,t_n) \to \forall x_1 \dots \forall x_n \varphi$, nicht wahr sind! (Was für bestimmte Objekte gilt, gilt nicht für beliebige: quod licet iovi...!)

**Distributionsregel**

$$\varphi \to \psi \vdash \exists x\varphi \to \exists x\psi$$
$$\varphi \to \psi \vdash \forall x\varphi \to \forall x\psi$$

*Beweis:*

| | | |
|---|---|---|
| (1) | $\varphi \to \psi$ | Voraussetzung |
| (2) | $\psi \to \exists x\psi$ | Substitutionssatz |
| | $\varphi \to \exists x\psi$ | tautologisch aus (1), (2) |
| | $\exists x\varphi \to \exists x\psi$ | mit $\exists$-Einführung (die Variablenbedingung ist erfüllt!) |

Den zweiten Teil beweist man analog (“dual”). □

Bereits früher hatten wir für eine Formel $\varphi = \varphi(x_1,\dots,x_n)$ mit $x_1,\dots,x_n$ als frei vorkommenden Variablen definiert:

$\forall x_1 \dots \forall x_n \varphi$ der **(universelle) Abschluß** von $\varphi$.

Dieser ist immer ein Satz (d. h. eine Formel ohne frei vorkommende Variable).

### 4.3.5 Abschlußsatz

*Ist $\varphi^{\forall}$ der universelle Abschluß von $\varphi$, so sind beide deduktionsgleich:*

$$\mathsf{T} \vdash \varphi \iff \mathsf{T} \vdash \varphi^{\forall}.$$

*Beweis:* Benutze die Generalisierungsregel bzw. den Substitutionssatz. □

Somit ist es gleichbedeutend, aus einer Theorie $\mathsf{T}$ eine Formel $\varphi(x,y,\ldots)$ oder ihren universellen Abschluß $\forall x \forall y\, \varphi(x,y,\ldots)$ zu beweisen:

$$\boxed{\forall x \forall y\, \varphi(x,y,\ldots) \text{ folgt aus } \mathsf{T} \text{ gdw } \varphi(x,y,\ldots) \text{ für "beliebige" } x,y \text{ aus } \mathsf{T} \text{ folgt.}}$$

## 4.4 Ersetzung und Umbenennung, Gleichheit

Wie in der Aussagenlogik gilt auch hier, daß der Wahrheitswert einer Formel sich nicht ändert, wenn man innerhalb der Formel eine Teilformel durch eine andere mit gleichem Wahrheitswert ersetzt. Hier zeigen wir das syntaktische Gegenstück: Ersetzt man innerhalb einer Formel $\varphi$ eine Teilformel durch eine äquivalente, so erhält man wiederum eine zu $\varphi$ äquivalente Formel. (Dabei braucht man die erwähnte Teilformel natürlich nicht überall zu ersetzen.)

### 4.4.1 Ersetzungstheorem

*Es gelte $\mathsf{T} \vdash \psi_i \leftrightarrow \psi_i'$ für $i = 1,\ldots n$, und die Formel $\varphi'$ entstehe aus $\varphi$, indem in $\varphi$ einige (oder alle) Teilformeln der Form $\psi_i$ durch $\psi_i'$ ersetzt werden. Dann gilt auch $\mathsf{T} \vdash \varphi \leftrightarrow \varphi'$.*

*Beweis* durch Induktion über den logischen Aufbau von $\varphi$. □

**Satz (Umbenennung von gebundenen Variablen)**

*Falls die Variable $y$ in $\psi$ nicht vorkommt:*

$$\vdash \exists x \psi \leftrightarrow \exists y \psi[y/x]$$

*Beweis:*

| | | |
|---|---|---|
| (1) | $\psi[y/x] \to \exists x\psi$ | Substitutionsaxiom |
| (2) | $\exists y\psi[y/x] \to \exists x\psi$ | $\exists$-Einführung in (1). Außerdem |
| (3) | $\psi \to \exists y\psi[y/x]$ | Substitutionsaxiom |
| (4) | $\exists x\psi \to \exists y\psi[y/x]$ | $\exists$-Einführung in (3) |

Dabei liegt in (3) eine Anwendung des Substitutionsaxioms auf die Formel $\psi[y/x]$ vor, wobei $\psi[y/x][x/y]$ wieder die Formel $\psi$ ergibt, da $x$ nicht in $\psi$ vorkommt. Dagegen wird in (2) nur benötigt, daß $x$ nicht frei in $\psi$ vorkommt. □

**Beispiel für die Notwendigkeit der Variablenbedingung:**

Es sei $\psi = \exists y(y \neq x)$. Dann ist nicht allgemeingültig (und damit auch nicht beweisbar):

$$\exists x\exists y(y \neq x) \leftrightarrow \exists y\exists y(y \neq y).$$

Tatsächlich ist hier $\psi[y/x][x/y] = \exists y(y \neq y) \neq \psi$.

Wir werden jetzt häufig der Einfachheit halber, soweit kein Mißverständnis zu befürchten ist, $\varphi(t_1,\dots,t_n)$ statt $\varphi[t_1/x_1,\dots,t_n/x_n]$ schreiben.

**Definition**

Die Formel $\varphi^*$ entsteht aus der Formel $\varphi$ durch **Umbenennung von gebundenen Variablen** gdw in $\varphi$ Teilformeln der Form $\exists x\psi(x)$ durch $\exists y\psi(y)$ ersetzt werden (wobei $y$ natürlich eine neue Variable ist, die in $\exists x\psi(x)$ nicht vorkommt).

**Satz (Gebundene Umbenennung)**

*Falls $\varphi^*$ aus der Formel $\varphi$ durch Umbenennung von gebundenen Variablen entsteht, so erhält man eine äquivalente Formel:*

$$\vdash \varphi \leftrightarrow \varphi^*$$

### 4.4.2 Eigenschaften des Gleichheitsprädikates

Zum Schluß behandeln wir noch Aussagen über die Gleichheit: In den Axiomen kommt an erster Stelle zwar nur die Reflexivität der Gleichheit vor, die übrigen

Eigenschaften sind aber in dem dritten Schema enthalten, in welchem die Relation $R$ auch das Gleichheitszeichen sein darf, so daß sich hier auch die Transitivität der Gleichheit versteckt. Es fehlt für eine Äquivalenzrelation somit nur noch die Symmetrie:

**Lemma**

$\vdash s = t \leftrightarrow t = s$

*Beweis:*

| | | |
|---|---|---|
| (1) | $x = x$ | Gleichheitsaxiom |
| (2) | $x = y \wedge x = x \wedge x = x \rightarrow y = x$ | Gleichheitsaxiom mit $Rx_1x_2$ als $x_1 = x_2$ |
| (3) | $x = y \rightarrow y = x$ | tautologisch aus (1), (2) |
| (4) | $s = t \rightarrow t = s$ | Substitutionsregel aus (3), ebenso |
| (5) | $t = s \rightarrow s = t$ | □ |

**Prinzip der Nichtunterscheidbarkeit**

(i) *$s'$ entstehe aus $s$ durch Ersetzung einiger (oder aller) Teilterme $t_i$ durch $t_i'$. Dann gilt:*

$$\vdash t_1 = t_1' \wedge \ldots \wedge t_n = t_n' \rightarrow s = s'.$$

(ii) *$t_i$ und $t_i'$ seien Terme, die in $\varphi$ substituierbar seien. Dann gilt:*

$$\vdash t_1 = t_1' \wedge \ldots \wedge t_n = t_n' \rightarrow (\varphi(t_1, \ldots, t_n) \leftrightarrow \varphi(t_1', \ldots, t_n')).$$

*Beweis:* Wegen der Substitutionsregel brauchen wir (i) und (ii) nur für Variable $x_i, y_i$ statt der Terme $t_i, t_i'$ zu beweisen. (i) beweist man durch Induktion über den Aufbau von $s$ (mit Hilfe der Gleichheitsaxiome), (ii) ähnlich durch Induktion über den Formelaufbau von $\varphi$. □

Das ist das wichtigste Ergebnis über die Gleichheit (in einfachster Form bereits in den Gleichheitsaxiomen enthalten): Im mathematischen Gebrauch (und nicht nur hier) ist die Gleichheitsbeziehung fast niemals die reine Identität (etwa im Sinne der syntaktischen Gleichheit), sondern stets nur eine Äquivalenzrelation und zusätzlich eine Kongruenzrelation bezüglich der Grundbegriffe der jeweiligen Theorie.

## 4.5 Pränexe Normalformen

Wir setzen unsere Untersuchungen fort mit weiteren Gesetzen über Quantoren, die u. a. dazu führen, pränexe Nomalformen zu erhalten, in welchen alle Quantoren am Anfang der Formel (als *Präfix*) stehen. In der Aussagenlogik konnten wir konjunktive und disjunktive Normalformen z. B. für Entscheidungsfragen benutzen; eine entsprechende Anwendung fehlt in der Prädikatenlogik, da diese unentscheidbar ist. Man kann aber pränexe Normalformen zur Klassifikation der Kompliziertheit benutzen (so wie man ein Polynom nach seinem Grad klassifizieren kann) und z. B. zeigen,

1. daß Formeln mit einem bestimmten Präfix noch entscheidbar, ab einem bestimmten Grad der Komplexität aber unentscheidbar sind,

2. Formeln aufgrund ihres Präfixes bestimmte modelltheoretische Eigenschaften haben.

3. Die Bedeutung der Quantorenreihenfolge am Anfang tritt außerdem offen zutage, wenn man etwa Eigenschaften der Stetigkeit mit derjenigen der gleichmäßigen Stetigkeit vergleicht, oder die Existenz eines größten Elementes: $\exists x \forall y (y < x)$ mit der Aussage, daß zu jedem Element ein größeres existiert: $\forall y \exists x (y < x)$.

Aus der Definition des $\forall$-Quantors mittels des $\exists$-Quantors ergibt sich:

### 4.5.1 Umformungen mit der Negation

$$\begin{array}{ll} (i) & \vdash \neg \forall x \varphi \leftrightarrow \exists x \neg \varphi \\ (ii) & \vdash \neg \exists x \varphi \leftrightarrow \forall x \neg \varphi \end{array}$$

### 4.5.2 Umformungen mit Konjunktion und Disjunktion

*Falls x in $\psi$ nicht frei vorkommt und Q den $\exists$- oder $\forall$-Quantor bezeichnet:*

$$\begin{array}{ll} (i) & \vdash Qx\varphi \vee \psi \leftrightarrow Qx(\varphi \vee \psi) \\ (ii) & \vdash \psi \vee Qx\varphi \leftrightarrow Qx(\psi \vee \varphi) \end{array}$$

*Ähnliche Aussagen gelten, wenn auf beiden Seiten* $\vee$ *durch* $\wedge$ *ersetzt wird.*

*Beweis:* Wir zeigen nur (i):

| | | |
|---|---|---|
| (1) | $\psi \to \varphi \vee \psi$ | Tautologie |
| (2) | $\varphi \vee \psi \to \exists x(\varphi \vee \psi)$ | Substitutionsaxiom |
| (3) | $\psi \to \exists x(\varphi \vee \psi)$ | tautologisch aus (1), (2) , ebenso |
| (4) | $\varphi \to \varphi \vee \psi$ | Tautologie |
| (5) | $\exists x\varphi \to \exists x(\varphi \vee \psi)$ | Distributionsregel, angewandt auf (4) |
| (6) | $\exists x\varphi \vee \psi \to \exists x(\varphi \vee \psi)$. | tautologisch aus (3), (5). Für die Umkehrung: |
| (7) | $\varphi \to \exists x\varphi$ | Substitutionsaxiom, daraus |
| (8) | $\varphi \vee \psi \to \exists x\varphi \vee \psi$ | als tautologische Folgerung |
| (9) | $\exists x(\varphi \vee \psi) \to \exists x\varphi \vee \psi$ | mit $\exists$-Einführung ($x$ nicht frei in $\varphi$) □ |

### 4.5.3 Umformungen mit der Implikation

*Falls x in $\psi$ nicht frei vorkommt:*

$$(i) \qquad \vdash (\forall x\varphi \to \psi) \leftrightarrow \exists x(\varphi \to \psi)$$
$$(ii) \qquad \vdash (\psi \to \forall x\varphi) \leftrightarrow \forall x(\psi \to \varphi)$$
$$(iii) \qquad \vdash (\exists x\varphi \to \psi) \leftrightarrow \forall x(\varphi \to \psi)$$
$$(iv) \qquad \vdash (\psi \to \exists x\varphi) \leftrightarrow \exists x(\psi \to \varphi)$$

Diese Ergebnisse kann man auf die früheren zurückführen, indem man die Äquivalenz $(\varphi \to \psi) \leftrightarrow (\neg\varphi \vee \psi)$ benutzt. Dadurch wird zugleich auch deutlich, warum in (i) und (iii) die Quantoren wechseln: die linke Seite des $\to$ erhält bei der Umformung ein $\neg$-Zeichen! □

Die folgenden Äquivalenzen kann man sich als eine Art Assoziativgesetz merken, wenn man den $\exists$-Quantor als verallgemeinertes $\vee$, den $\forall$-Quantor als verallgemeinertes $\wedge$ auffaßt:

**Lemma**

$$(i) \qquad \vdash \exists x(\varphi \vee \psi) \leftrightarrow \exists x\varphi \vee \exists x\psi$$
$$(ii) \qquad \vdash \forall x\varphi \wedge \forall x\psi \leftrightarrow \forall x(\varphi \wedge \psi)$$

Die folgenden Ergebnisse gelten auch ohne Variablenbedingung, dafür i. a. aber nur in jeweils einer Richtung:

**Lemma**

$$\vdash \forall x\varphi \to \exists x\varphi$$
$$\vdash \exists x(\varphi \wedge \psi) \to \exists x\varphi \wedge \exists x\psi$$
$$\vdash (\forall x\varphi \vee \forall x\psi) \to \forall x(\varphi \vee \psi)$$
$$\vdash (\exists x\varphi \to \exists x\psi) \to \exists x(\varphi \to \psi)$$
$$\vdash \forall x(\varphi \to \psi) \to (\forall x\varphi \to \forall x\psi)$$

**Definition (pränex)**

$\varphi$ heißt **offen** gdw $\varphi$ keine Quantoren enthält, $\varphi$ heißt **pränex** gdw $\varphi$ offen ist oder von der Form

$$Q_1 x_1 \dots Q_n x_n \, \psi$$

mit $\psi$ offen, $Q_i$ ein Quantor ($\exists$ oder $\forall$). Dabei heißt der Quantorenblock $Q_1 x_1 \dots Q_n x_n$ das **Präfix** von $\varphi$, $\psi$ die **Matrix** von $\varphi$.

**Satz über die pränexe Normalform**

*Zu jeder Formel $\varphi$ (ausgedrückt mit $\neg, \vee, \wedge, \to, \leftrightarrow, \exists, \forall$) existiert eine pränexe Formel $\varphi^p$ mit*

$$\vdash \varphi \leftrightarrow \varphi^p.$$

*$\varphi^p$ heißt (eine)* ***pränexe Normalform*** *von $\varphi$.*

Den *Beweis* führt man durch Induktion über den Aufbau der Formel $\varphi$. Wir wollen stattdessen skizzieren, wie man eine pränexe Normalform konkret erhält: Gegeben sei eine Formel $\varphi$, von der wir annehmen können, daß sie außer den Quantoren nur die BOOLEschen Operationen $\neg, \vee, \wedge$ enthält.

(1) Mit Hilfe von 4.5.1 und den entsprechenden aussagenlogischen Umformungen bringe man in $\varphi$ das $\neg$-Zeichen so weit nach innen, bis es höchstens vor atomaren Formeln steht; das Ergebnis ist eine äquivalente Formel $\varphi^n$ *(Negations-Normalform)*.

(2) Bringe den in $\varphi^n$ am weitesten links stehenden Quantor (u. U. nach geeigneter Umbenennung) an den Anfang der Formel und erstrecke seinen

Wirkungsbereich auf die ganze Formel; benutze dazu die folgenden Äquivalenzen (aus den vorhergehenden Lemmata):

$$
\begin{array}{lll}
\forall v\varphi(v) \wedge \theta & \leftrightarrow \forall v(\varphi(v) \wedge \theta) & \text{falls } v \text{ nicht in } \theta \\
& \leftrightarrow \forall v(\varphi(v) \wedge \theta'(v)) & \text{falls } \theta = \forall v\theta'(v) \\
\forall v\varphi(v) \vee \theta & \leftrightarrow \forall v(\varphi(v) \vee \theta) & \text{falls } v \text{ nicht in } \theta \\
& \leftrightarrow \forall z(\varphi(z) \vee \theta) & \text{mit geeignetem } z
\end{array}
$$

Die Umbenennung der Variablen $v$ in eine geeignete neue Variable $z$ kann man natürlich immer vornehmen; die vorhergehenden Umformungen sind - soweit möglich - jedoch vorzuziehen, um die neue Formel nicht unnötig kompliziert zu machen.

Auf ähnliche Weise verfahre man im Falle eines $\exists$-Quantors. Damit erhalten wir eine zu $\varphi^n$ äquivalente Formel $\varphi'$ der Form $Qv\delta$. Bringt man dann in gleicher Weise die in $\delta$ vorkommenden Quantoren “an den Anfang”, so erhält man als Endergebnis eine pränexe Formel $\varphi^p$, die zu $\varphi$ äquivalent ist. □

**Bemerkung:**

Wie im Falle der konjunktiven bzw. disjunktiven Normalform ist eine pränexe Normalform niemals eindeutig bestimmt (trotzdem spricht man häufig von *der* pränexen Normalform, da zumindest alle untereinander äquivalent sind).

### 4.5.4 Das Dualitätsprinzip für die Prädikatenlogik

Wir setzen hier voraus, daß die Formeln (außer den Quantoren) wieder nur die BOOLEschen Operationen $\neg, \vee, \wedge$ enthalten. Dann definiert man in Verallgemeinerung des aussagenlogischen Falles:

$D(\varphi)$ : vertausche in $\varphi$ $\vee$ mit $\wedge$ und $\exists$ mit $\forall$ : **duale Formel** zu $\varphi$,

$N(\varphi)$ : setze vor jede atomare Formel in $\varphi$ ein $\neg$-Zeichen,

$N^*(\varphi)$ : setze vor jede atomare Formel in $\varphi$ ein $\neg$-Zeichen, falls dort keines steht, sonst streiche es!

**Beispiel:**

Es sei $\varphi = \forall x(P(x) \rightarrow \exists yR(y,x))$, d. h. nach Elimination von $\rightarrow$:
$\varphi \leftrightarrow \forall x(\neg P(x) \vee \exists yR(y,x))$.

Dann ist

$$\begin{aligned} D(\varphi) &= \exists x(\neg P(x) \wedge \forall y R(y,x)) \\ N(\varphi) &= \forall x(\neg\neg P(x) \vee \exists y \neg R(y,x)) \\ N^*(\varphi) &= \forall x(P(x) \vee \exists y \neg R(y,x)) \\ DN^*(\varphi) &= \exists x(P(x) \wedge \forall y \neg R(y,x)) \end{aligned}$$

Ähnlich wie in der Aussagenlogik erhält man (hier in syntaktischer Form):

**Satz**

| | | |
|---|---|---|
| a) | $\vdash \neg\varphi \leftrightarrow D(N^*(\varphi))$ | **Bildung der Negation** |
| b) | $\vdash (\varphi \leftrightarrow \psi) \iff \vdash (D\varphi \leftrightarrow D\psi)$ | **Dualitätssatz** |

**Beschränkte Quantoren**

Man kann diese Ergebnisse auch auf beschränkte Quantoren erweitern: Ist

$$\begin{aligned} \varphi &= \forall x(P(x) \to \psi(x)), \text{ d. h.} \\ &\leftrightarrow \forall x(\neg P(x) \vee \psi(x)), \text{ so ist} \\ \neg\varphi &\leftrightarrow \exists x(P(x) \wedge \neg\psi(x)). \end{aligned}$$

d. h. die **beschränkten Quantoren**

$$\forall x(P(x) \to \ldots) \quad \text{und} \quad \exists x(P(x) \wedge \ldots)$$

sind dual zueinander.

Besonders häufig kommt ein derartiger Quantor in der Analysis vor als

$$\forall \varepsilon > 0\ \psi :\leftrightarrow \forall \varepsilon\,(\varepsilon > 0 \to \psi)$$

Mit Hilfe des obigen Satzes lassen sich also Aussagen der folgenden Form leicht negieren:

$$\neg\forall x \forall \varepsilon > 0 \exists \delta > 0\, \psi \leftrightarrow \exists x \exists \varepsilon > 0 \forall \delta > 0 \neg\psi.$$

## 4.6 Das Deduktionstheorem

Beim Beweisen steht man oft vor der folgenden Aufgabe: In einer Theorie $\mathsf{T}$ soll eine Implikation der Form $\varphi \to \psi$ gezeigt werden. Dazu benutzt man das folgende Argument:

"Vorausgesetzt sei $\varphi$. Dann ... (hier folgt der Beweis mit den zusätzlichen Annahmen aus $\mathsf{T}$) ... : $\psi$. Also haben wir $\mathsf{T} \vdash \varphi \to \psi$ gezeigt."

Das Deduktionstheorem rechtfertigt dieses Argument, falls eine wichtige Einschränkung beachtet wird:

**Deduktionstheorem**

$\mathsf{T}[\sigma]$ *sei die Theorie* $\mathsf{T}$ *mit* $\sigma$ *als zusätzlichem Axiom und* $\sigma$ *ein Satz. Dann gilt:*

$$\mathsf{T} \vdash \sigma \to \psi \iff \mathsf{T}[\sigma] \vdash \psi$$

Der *Beweis von* "$\Rightarrow$" ist trivial und macht nur vom modus ponens Gebrauch:

Es gelte $\mathsf{T} \vdash \sigma \to \psi$, dann offensichtlich also auch $\mathsf{T}[\sigma] \vdash \sigma \to \psi$. Da ferner $\mathsf{T}[\sigma] \vdash \sigma$, so mit modus ponens $\mathsf{T}[\sigma] \vdash \psi$.

Der Teil "$\Leftarrow$" ist interessanter und auch weniger trivial: Beim Beweis von $\psi$ aus $\mathsf{T}[\sigma]$ könnte das Zusatzaxiom $\sigma$ an mehreren Stellen benutzt worden sein! Wir werden zeigen, daß man aus einem Beweis von $\psi$ aus $\mathsf{T}[\sigma]$ (mit leichten Modifikationen) einen Beweis von $\sigma \to \psi$ aus $\mathsf{T}$ erhält, indem man im ersten Beweis jede Beweiszeile $\delta$ durch $\sigma \to \delta$ ersetzt, wobei das Zusatzaxiom $\sigma$ in die beweisbare Formel $\sigma \to \sigma$ übergeht und damit entbehrlich wird. Aufpassen muß man allerdings im Falle der $\exists$-Einführungsregel, daß die Variablenbedingung nicht verletzt wird; die Voraussetzung, daß $\sigma$ ein Satz ist, wird hier benötigt (könnte hier aber auch entsprechend abgeschwächt werden).

*Beweis* von "$\Leftarrow$": Wir zeigen

$$\mathsf{T}[\sigma] \vdash \psi \Longrightarrow \mathsf{T} \vdash \sigma \to \psi$$

durch Induktion über die Länge eines Beweises (d. h. Anzahl der Formeln im Beweis) von $\psi$:

a) $\psi$ ist Axiom von $\mathsf{T}[\sigma]$: dann ist entweder

   (i) $\psi$ Axiom von $\mathsf{T}$ und damit $\mathsf{T} \vdash \psi$, also auch $\mathsf{T} \vdash \sigma \to \psi$, oder

   (ii) $\psi = \sigma$ und damit wegen $\mathsf{T} \vdash \sigma \to \sigma$ auch $\mathsf{T} \vdash \sigma \to \psi$.

b) $\psi$ ist logisches Axiom. Dann gilt $\mathsf{T} \vdash \psi$, also auch (als tautologische Folgerung) $\mathsf{T} \vdash \sigma \to \psi$.

Damit haben wir den Induktionsanfang behandelt. Der Induktionsschluß ($\psi$ ist Folgerung mittels einer Regel) läßt sich auch in 2 Fälle aufgliedern:

c) Es gibt Formeln $\psi_1$ bzw. $\psi_1$ und $\psi_2$, so daß $\psi$ aus diesen mittels einer aussagenlogischen Regel folgt (und damit tautologisch folgt), wobei $\mathsf{T}[\sigma] \vdash \psi_1$ und $\mathsf{T}[\sigma] \vdash \psi_2$. Nach Induktionsvoraussetzung gilt ferner $\mathsf{T} \vdash \sigma \to \psi_i$ für $i = 1,2$ und damit als tautologische Folgerung auch $\mathsf{T} \vdash \sigma \to \psi$.

d) $\psi$ ist Folgerung aus $\psi_1$ mittels der Regel der $\exists$-Einführung, also
$\psi_1 = \delta_1 \to \delta_2$ und $\psi = \exists x \delta_1 \to \delta_2$ für Formeln $\delta_1, \delta_2$, wobei $x$ in $\delta_2$ nicht frei vorkommt. Dann:

| | | |
|---|---|---|
| (1) | $\mathsf{T} \vdash \sigma \to (\delta_1 \to \delta_2)$ | nach Induktionsvoraussetzung |
| (2) | $\mathsf{T} \vdash \delta_1 \to (\sigma \to \delta_2)$ | tautologisch (aus (1) |
| (3) | $\mathsf{T} \vdash \exists x \delta_1 \to (\sigma \to \delta_2)$ | mit $\exists$-Einführung |
| (4) | $\mathsf{T} \vdash \sigma \to (\exists x \delta_1 \to \delta_2)$ | tautologisch aus (3), also |
| (5) | $\mathsf{T} \vdash \sigma \to \psi$ | |

Dabei war die $\exists$-Einführung in (3) erlaubt, denn $x$ kommt nach Voraussetzung in $\delta_2$ nicht frei vor, aber auch nicht frei in $\sigma$ (da nach Annahme $\sigma$ ein Satz ist), also kommt $x$ in $\sigma \to \delta_2$ nicht frei vor. $\square$

**Korollar**

$\mathsf{T}$ *sei eine Theorie der Sprache* $\mathcal{L}$, $\gamma_1, \ldots, \gamma_n$ seien $\mathcal{L}$ *-Sätze und* $\psi$ *eine* $\mathcal{L}$*-Formel. Dann gilt:*

$$\mathsf{T} \vdash \gamma_1 \wedge \ldots \wedge \gamma_n \to \psi \iff \mathsf{T}[\gamma_1, \ldots, \gamma_n] \vdash \psi$$

**Bemerkungen**

1. Daß das Deduktionstheorem nicht für beliebige Formeln gilt, zeigt das Beispiel $\varphi = (x = y)$, wobei $\mathcal{A} \models \forall x \forall y\, \varphi \iff |\mathcal{A}|$ hat genau ein Element:

   Es ist $\mathsf{T}[x = y] \vdash \forall x \forall y (x = y)$, aber $\mathsf{T} \nvdash x = y \to \forall x \forall y (x = y)$, wenn $\mathsf{T}$ ein Modell mit mindestens zwei Elementen hat.

2. Aus dem Beweis ergibt sich, daß $\sigma$ auch eine Formel sein kann, vorausgesetzt, daß beim Beweis von $\psi$ aus $\mathsf{T}[\sigma]$ die $\exists$-Einführungsregel nicht für eine Variable benutzt wurde, die in $\sigma$ frei vorkommt. Ansonsten läßt sich das Deduktionstheorem für eine *Formel* $\varphi$ auch in der Form

$$\mathsf{T} \vdash \varphi^{\forall} \to \psi \iff \mathsf{T}[\varphi] \vdash \psi$$

umformulieren, wobei $\varphi^{\forall}$ der universelle Abschluß von $\varphi$ ist.

3. Falls $\mathsf{T} \vdash \psi$, so existieren nach dem Endlichkeitssatz für die Beweisbarkeit endlich viele $\varphi_1, \ldots, \varphi_n$ in $\mathsf{T}$ mit $\varphi_1, \ldots, \varphi_n \vdash \psi$; sind die Axiome $\varphi_1, \ldots, \varphi_n$ Sätze (man kann sie anderenfalls wegen des Abschlußsatzes durch ihre universellen Abschlüsse ersetzen), so gilt also $\vdash \varphi_1 \wedge \ldots \wedge \varphi_n \to \psi$.

Damit läßt sich die Beweisbarkeit in einer Theorie (prinzipiell) auf logische Beweisbarkeit zurückführen; ist ein konkreter Beweis von $\psi$ aus $\mathsf{T}$ gegeben, so lassen sich aus diesem Beweis die Sätze $\varphi_1, \ldots, \varphi_n$ in $\mathsf{T}$ mit $\varphi_1 \wedge \ldots \wedge \varphi_n \vdash \psi$ auch konkret angeben, nicht notwendig aber, wenn man nur von der Beweis<u>bar</u>keit ausgeht!

## 4.7 Erweiterungen von Theorien, Widerspruchsfreiheit

Theorien erweitert man oft durch Hinzunahme neuer Axiome (z. B. des Kommutativgesetzes zu den Gruppenaxiomen oder der Negation des Parallelenaxioms zur "Absoluten" Geometrie), manchmal wird dabei auch noch die Sprache erweitert (Gruppentheorie $\to$ Körpertheorie $\to$ Theorie der angeordneten Körper).

### 4.7.1 Erweiterungen und Expansionen

a) Eine Sprache $\mathcal{L}'$ heißt **Erweiterung** der Sprache $\mathcal{L}$ ($\mathcal{L} \subseteq \mathcal{L}'$) gdw. jedes (nicht-logische) Symbol von $\mathcal{L}$ auch Symbol von $\mathcal{L}'$ ist (also jede $\mathcal{L}$-Formel auch eine $\mathcal{L}'$-Formel ist).

b) Eine Theorie $\mathsf{T}'$ der Sprache $\mathcal{L}'$ heißt **Erweiterung** der Theorie $\mathsf{T}$ der Sprache $\mathcal{L}$ ($\mathsf{T} \subseteq \mathsf{T}'$) gdw

   (i) $\mathcal{L}'$ Erweiterung der Sprache $\mathcal{L}$ ist und

   (ii) für alle Formeln $\varphi$ der Sprache $\mathcal{L}$: $\mathsf{T} \vdash \varphi \Rightarrow \mathsf{T}' \vdash \varphi$.

c) Gilt in (ii) sogar die Äquivalenz, also (zusätzlich zu (i))

   (iii) für alle Formeln $\varphi$ der Sprache $\mathcal{L}$: $\mathsf{T} \vdash \varphi \iff \mathsf{T}' \vdash \varphi$,

   so heißt $\mathsf{T}'$ **konservative Erweiterung** von $\mathsf{T}$.

Häufig erweitert man eine Theorie $\mathsf{T}$ durch Definitionen: Führt man für eine Formel $\varphi(\vec{x})$ ein neues Relationszeichen $R$ ein, so bedeutet dies, daß man $\mathsf{T}$ zu einer Theorie $\mathsf{T}'$ erweitert, die als neues Axiom erhält:

$$R(\vec{x}) \leftrightarrow \varphi(\vec{x}).$$

Ebenso bei der Einführung neuer Funktionszeichen: Hier muß man zunächst nachweisen, daß

$$\mathsf{T} \vdash \forall \vec{x} \exists ! y \, \varphi(\vec{x}, y),$$

worauf man durch

$$f(\vec{x}) = y \leftrightarrow \varphi(\vec{x}, y)$$

die Theorie $\mathsf{T}$ um das neue Funktionszeichen $f$ mit obigem Axiom erweitert. In beiden Fällen ist die neue Theorie eine konservative Erweiterung der alten Theorie $\mathsf{T}$.

d) Falls $\mathsf{T} \subseteq \mathsf{T}'$ und $\mathsf{T}' \subseteq \mathsf{T}$ , so heißt $\mathsf{T}$ **äquivalent** zu $\mathsf{T}'$ (und dann wird oft $\mathsf{T}$ mit $\mathsf{T}'$ identifiziert, da sie sich in diesem Fall (möglicherweise) nur durch ihre Axiomatisierung, nicht aber in ihren Konsequenzen voneinander unterscheiden).

Insbesondere unterscheiden wir nicht zwischen einer Theorie $\mathsf{T}$ und ihrer Folgerungsmenge $\mathsf{C}(\mathsf{T}) := \{\sigma \mid \mathsf{T} \vdash \sigma\}$. Mit Hilfe des Vollständigkeitssatzes werden wir später zeigen, daß Theorien genau dann äquivalent sind, wenn sie dieselben Modelle haben.

Beachte, daß äquivalente Theorien in derselben Sprache formuliert sind und daß eine konservative Erweiterung $\mathsf{T}'$ von $\mathsf{T}$ zwar nur dieselben Aussagen der (kleineren) Sprache von $\mathsf{T}$ abzuleiten gestattet, darüber hinaus jedoch möglicherweise weitere Aussagen über die erweiterte Sprache zuläßt.

e) Ist $\mathcal{L} \subseteq \mathcal{L}'$, $\mathcal{A}$ eine $\mathcal{L}$-Struktur und $\mathcal{A}'$ eine $\mathcal{L}'$-Struktur, so heißt

   $\mathcal{A}'$ **Expansion** von $\mathcal{A} :\iff |\mathcal{A}| = |\mathcal{A}'|$ und $s^{\mathcal{A}} = s^{\mathcal{A}'}$ für alle nichtlogischen Symbole $s$ von $\mathcal{L}$.

   Man schreibt dann auch $\mathcal{A} = \mathcal{A}' \upharpoonright \mathcal{L}$ und nennt umgekehrt $\mathcal{A}$ auch die **Einschränkung** oder das **Redukt** von $\mathcal{A}'$ auf die Sprache $\mathcal{L}$.

So ist z. B. ein Körper $(K,+,\cdot,0,1)$ eine Expansion der additiven Gruppe $(K,+,0)$ des Körpers.

Eine Erweiterung einer Theorie $\mathsf{T}$ der Sprache $\mathcal{L}$ zu einer Theorie $\mathsf{T}'$ der Sprache $\mathcal{L}'$ kann dadurch erfolgen, daß nur die Sprache erweitert wird, aber keine weiteren nicht-logischen Axiome hinzukommen ($\mathsf{T}'$ heißt dann **rein sprachliche Erweiterung** von $\mathsf{T}$) oder es können (wie beim Übergang von der Gruppen- zur Körpertheorie) noch weitere nicht-logische Axiome hinzukommen. Wir behandeln zunächst den ersten Fall:

### 4.7.2 Satz über rein sprachliche Erweiterungen von Theorien

$\mathsf{T}$ *und* $\mathsf{T}'$ *seien Theorien der Sprache* $\mathcal{L}$ *bzw.* $\mathcal{L}'$*, wobei* $\mathcal{L} \subseteq \mathcal{L}'$*. Haben* $\mathsf{T}$ *und* $\mathsf{T}'$ *dieselben Axiome, so ist* $\mathsf{T}'$ *konservative Erweiterung von* $\mathsf{T}$*:*

$$\textit{für alle Formeln } \varphi \textit{ von } \mathcal{L} \textit{ gilt: } \mathsf{T} \vdash \varphi \iff \mathsf{T}' \vdash \varphi.$$

*Vorbemerkung:* Der Teil der Behauptung von links nach rechts ist trivial (da jedes Axiom von $\mathsf{T}$ auch Axiom von $\mathsf{T}'$ ist, ist jeder Beweis von $\varphi$ aus $\mathsf{T}$ auch ein Beweis von $\varphi$ aus $\mathsf{T}'$), aber die Umkehrung ist keineswegs trivial, da ein Beweis von $\varphi$ aus $\mathsf{T}'$ logische Axiome (bzw. Regeln) für Formeln der erweiterten Sprache benutzen kann (auch wenn die zusätzlichen Symbole der Sprache $\mathcal{L}'$ in $\varphi$ selbst nicht mehr vorkommen). Ein semantischer Beweis (mit $\models$ statt $\vdash$ ) ist jedoch einfach: ist $\mathcal{A}' \models \mathsf{T}'$ , so ist die Einschränkung $\mathcal{A} = \mathcal{A}' \upharpoonright \mathcal{L}$ ein Modell von $\mathsf{T}$.

*Beweis des Satzes:* In einer Formel $\psi$ der (größeren) Sprache $\mathcal{L}'$ streichen wir die nicht zu $\mathcal{L}$ gehörigen Symbole und ersetzen sie in harmloser Weise mittels einer Variablen $y$ : Es sei $\psi_y$ die Formel der Sprache $\mathcal{L}$ , die aus $\psi$ entsteht, indem in $\psi$

| | | | |
|---|---|---|---|
| $Rt_1 \ldots t_n$ | (falls $R$ nicht in $\mathcal{L}$) durch | $y = y$, | |
| $Ft_1 \ldots t_m$ | (falls $F$ nicht in $\mathcal{L}$) durch | $y$, | |
| $c$ | (falls $c$ nicht in $\mathcal{L}$) durch | $y$ | ersetzt wird. |

Wir zeigen nun für alle Formeln $\psi$ der Sprache $\mathcal{L}'$:

(*) für alle Variablen $y$ bis auf endlich-viele Ausnahmen gilt $\mathsf{T}' \vdash \psi \Rightarrow \mathsf{T} \vdash \psi_y$.

Da für Formeln $\psi$ in $\mathcal{L}$ offenbar $\psi_y = \psi$ ist, so folgt hieraus dann die Behauptung. Die Aussage (*) beweisen wir durch Induktion über die Länge eines Beweises von $\mathsf{T}' \vdash \psi$:

1. $\psi$ ist nicht-logisches Axiom von $\mathsf{T}'$. Dann ist $\psi$ auch (nicht-logisches) Axiom von $\mathsf{T}$, also $\mathsf{T} \vdash \psi$ und dann auch $\mathsf{T} \vdash \psi_y$ für alle Variablen $y$ (da hier $\psi_y = \psi$).

2. $\psi$ ist logisches Axiom von $\mathsf{T}'$.

   2.1 $\psi = \neg\theta \vee \theta$.
   Dann ist $\psi_y = \neg\theta_y \vee \theta_y$ für jedes $y$ und $\mathsf{T} \vdash \psi_y$ für alle Variablen $y$.

   2.2 $\psi = \theta[t/x] \to \exists x\theta$.
   Dann ist $\psi_y = \theta_y[t_0/x] \to \exists x\theta_y$ für alle $y$, die in $\theta$ nicht vorkommen und für einen Term $t_0$ von $\mathcal{L}$. Somit $\mathsf{T} \vdash \psi_y$ für alle Variablen $y$ bis auf endlich-viele.

   2.3 $\psi = (x = x)$. Dann ist $\psi_y = (x = x)$.

   2.4 $\psi$ ist eines der Gleichheitsaxiome.
   Dann ist $\psi_y = \psi$ oder von der Form $\theta \to y = y$, jedenfalls $\mathsf{T} \vdash \psi_y$ für alle Variablen $y$.

3. (Induktionsschritt):
   $\psi$ ist Konklusion einer Regel mit den Prämissen $\psi_1 \dots \psi_n \quad (n = 1,2)$, wobei nach Induktionsvoraussetzung $\mathsf{T} \vdash (\psi_i)_y$ für $i \leq n$ und alle Variablen $y$ bis auf endlich-viele gilt.

   3.1 Im Falle einer aussagenlogischen Regel ist $\psi$ tautologische Folgerung aus $\psi_1, \dots, \psi_n$, dann ist aber auch $\psi_y$ tautologische Folgerung aus $(\psi_1)_y, \dots, (\psi_n)_y$ und somit $\mathsf{T} \vdash \psi_y$.

   3.2 Im Falle der vorderen Partikularisierung ist $n = 1$ und $\psi = \exists x\varphi \to \theta$ und $\psi_1 = \exists x\varphi \to \theta$ mit $x$ nicht frei in $\theta$.
   Dann ist $\psi_y = \exists x\varphi_y \to \theta_y$ und $(\psi_1)_y = \varphi_y \to \theta_y$ mit $x$ nicht frei in $\theta_y$ für $y \neq x$. Also $\mathsf{T} \vdash \psi_y$ für jedes $y \neq x$ mit $\mathsf{T} \vdash (\psi_1)_y$ und damit $\mathsf{T} \vdash \psi_y$ für alle $y$ bis auf endlich-viele. $\square$

Das folgende Korollar haben wir einfacher schon in der modelltheoretischen Fassung (3.1.1) erhalten:

**Korollar**

*$\mathsf{T}$ sei eine Theorie der Sprache $\mathcal{L}$, $c$ eine neue Konstante, die in $\mathcal{L}$ nicht vorkommt. Dann gilt:*

$$\mathsf{T} \vdash \varphi[c/x] \iff \mathsf{T} \vdash \forall x\varphi. \quad \square$$

Hier haben wir erneut ein bekanntes Beweisprinzip (vgl. Satz über die Generalisierung) erhalten:

| |
|---|
| Um aus den Axiomen einer Theorie $\mathsf{T}$ eine Aussage $\forall x\varphi(x)$ zu beweisen, genügt es, die Aussage $\varphi(c)$ für eine geeignete Konstante $c$ (ein "beliebiges, aber festes" $c$) herzuleiten. |

Im Rahmen der Aussagenlogik hatten wir bereits früher die Widerspruchsfreiheit einer Aussagenmenge formuliert, entsprechend definieren wir im Falle der Prädikatenlogik:

### 4.7.3 Widerspruchsfrei/widerspruchsvoll

$\mathsf{T}$ **konsistent** : $\iff$ es gibt einen Satz $\sigma$ der Sprache von $\mathsf{T}$ mit: $\mathsf{T} \nvdash \sigma$,

$\mathsf{T}$ **inkonsistent** : $\iff$ für alle Sätze $\sigma$ der Sprache von $\mathsf{T}$ gilt: $\mathsf{T} \vdash \sigma$.

Diese Definition (statt *konsistent* sagt man auch *widerspruchsfrei*, statt *inkonsistent* auch *widerspruchsvoll*) hat den Vorzug, daß sie auch für Sprachen ohne das $\neg$-Zeichen sinnvoll ist; hier ergibt sich für unsere Sprachen die Äquivalenz mit der erwarteten Charakterisierung:

### Lemma

$$\mathsf{T} \textit{ konsistent} \iff \textit{es gibt keinen Satz } \sigma \textit{ mit } \mathsf{T} \vdash \sigma \textit{ und } \mathsf{T} \vdash \neg\sigma.$$

*Beweis* von "$\Rightarrow$": Falls $\mathsf{T} \vdash \sigma$ und $\mathsf{T} \vdash \neg\sigma$ für einen Satz $\sigma$, so nach dem Tautologiesatz $\mathsf{T} \vdash \sigma \wedge \neg\sigma$ und somit $\mathsf{T} \vdash \psi$ für alle Sätze $\psi$.
"$\Leftarrow$": Wähle für $\sigma$ den Satz $\forall x(x = x)$; dann gilt $\mathsf{T} \vdash \sigma$, also nach Vor. $\mathsf{T} \nvdash \neg\sigma$, und damit ist mindestens ein Satz aus $\mathsf{T}$ nicht beweisbar. $\square$

### Bemerkungen

1. *Ist* $\mathsf{T}'$ *Erweiterung von* $\mathsf{T}$ *und ist* $\mathsf{T}'$ *konsistent, so auch* $\mathsf{T}$.

2. *Ist* $\mathsf{T}'$ *konservative Erweiterung von* $\mathsf{T}$*, so gilt:*
   $\mathsf{T}$ *konsistent gdw* $\mathsf{T}'$ *konsistent.*

Nimmt man zu einer Theorie $\mathsf{T}$ eine Menge $\Gamma$ Formeln (als zusätzliche Axiome) hinzu, so ist die entstehende Erweiterung $\mathsf{T}[\Gamma]$ widerspruchsvoll, wenn endlich-viele der neuen Axiome einen Widerspruch zur Theorie $\mathsf{T}$ ergeben. Da man hier endlich-viele Axiome durch ihre Konjunktion ersetzen kann, genügt der folgende Satz:

### 4.7.4 Beweisbarkeit und Widerspruchsfreiheit

*Ist $\varphi^*$ der universelle Abschluß von $\varphi$ , so gilt:*

$$\mathsf{T} \vdash \varphi \iff \mathsf{T}[\neg\varphi^*] \textit{ widerspruchsvoll,}$$
$$\mathsf{T} \nvdash \varphi \iff \mathsf{T}[\neg\varphi^*] \textit{ widerspruchsfrei.}$$

*Beweis* von "$\Rightarrow$": Falls $\mathsf{T} \vdash \varphi$, so auch $\mathsf{T} \vdash \varphi^*$ und damit erst recht $\mathsf{T}[\neg\varphi^*] \vdash \varphi^*$. Andererseits gilt aber natürlich auch $\mathsf{T}[\neg\varphi^*] \vdash \neg\varphi^*$. Somit ist $\mathsf{T}[\neg\varphi^*]$ widerspruchsvoll.
"$\Leftarrow$": Es sei $\mathsf{T}[\neg\varphi^*]$ widerspruchsvoll. Dann ist alles beweisbar, insbesondere $\mathsf{T}[\neg\varphi^*] \vdash \varphi^*$. Nach dem Deduktionstheorem: $\mathsf{T} \vdash \neg\varphi^* \to \varphi^*$, also $\mathsf{T} \vdash \varphi^*$ und damit auch $\mathsf{T} \vdash \varphi$. □

Damit kann die Beweisbarkeit eines Satzes auf die Widerspruchsfreiheit einer geeignet erweiterten Theorie zurückgeführt werden:

$$\boxed{\mathsf{T} \nvdash \sigma \iff \mathsf{T}[\neg\sigma] \text{ ist widerspruchsfrei}}$$

Schon im Rahmen der Aussagenlogik haben wir bemerkt, daß der Begriff der Folgerung sich auf den entsprechenden Begriff der Erfüllbarkeit zurückführen läßt:

$$\boxed{\mathsf{T} \not\models \sigma \iff \mathsf{T}[\neg\sigma] \text{ ist erfüllbar}}$$

Wie im Falle der Aussagenlogik können wir den Vollständigkeitssatzes damit auf den **Modellexistenz-Satz** zurückführen:

*Jede widerspruchsfreie Theorie $\mathsf{T}$ ist erfüllbar.*

Tatsächlich ist diese Aussage äquivalent zum Vollständigkeitssatz; beide stellen einen Zusammenhang zwischen einem syntaktischen Begriff (beweisbar, widerspruchsfrei) und einem semantischen Begriff (folgt aus, erfüllbar) her. Der Modellexistenzsatz erscheint hier nur in seiner einfachsten Form; in der *Modelltheorie* werden allgemeinere Fassungen dieses Satzes behandelt, in denen aus syntaktischen Eigenschaften einer formalen Theorie Aussagen über semantische Eigenschaften möglicher Modelle gemacht werden.

## 4.8 Termmodelle

$\mathsf{T}$ sei eine Theorie der Sprache $\mathcal{L}$. Um ein Modell $\mathcal{A}$ von $\mathsf{T}$ zu finden, muß man wenigstens den Individuenkonstanten der Sprache Werte zuordnen; unter günstigen Voraussetzungen erhält man dann durch geeignete Festlegungen der Interpretation der Relations- und Funktionssymbole bereits ein Modell. So findet man z. B. im Falle der Gruppentheorie (und ähnlich im Falle der Theorie der Körper) das einfachste Modell:

> Wähle ein Element $e$ als (Interpretation des) neutrales Element(es) und setze $e \circ e = e$, $e^{-1} = e : G = (\{e\}, \circ, ^{-1}, e)$ ist eine Gruppe!

Diesen Ansatz wollen wir weiterverfolgen:

### 4.8.1 Kanonische Struktur einer Theorie, Termstruktur

Die Elemente von eines Modell $\mathcal{A}$ einer Theorie $\mathsf{T}$ müssen mindestens Interpretationen der konstanten Terme als Elemente enthalten. Daher definieren wir:

$$K := \{t \mid t \textit{ konstanter (d. h. variablenfreier) } \mathcal{L}\textit{-Term}\}.$$

Zusätzlich ist zu beachten ist, daß aus den Axiomen der Theorie $\mathsf{T}$ folgen kann, daß verschiedene Terme gleichgesetzt werden sollen (z. B. $0+1 = 1+(0\cdot 1)$ in der Körpertheorie). Daher werden wir Terme identifizieren, wenn die Theorie es verlangt:

$$s \sim t :\iff \mathsf{T} \vdash s = t \qquad \text{für } s,t \in K.$$

Diese Relation ist (wegen der Gleichheitsaxiome!) eine Äquivalenzrelation auf $K$,

$$\begin{aligned} \overline{s} &:= \{t \in K \mid s \sim t\} && \text{zugehörige Äquivalenzklasse.} \\ A &:= \{\overline{s} \mid s \in K\} && \text{Menge der Äquivalenzklassen} \end{aligned}$$

wird der Individuenbereich der gesuchten Struktur sein, in dem $t^{\mathcal{A}} = \overline{t}$ für alle $t \in K$ gelten wird. Dazu legen wir die Interpretation der weiteren nicht-logischen Zeichen wie folgt fest:

$$\begin{array}{rcl} c^{\mathcal{A}} & = & \overline{c} \\ F^{\mathcal{A}}(\overline{t_1}, \dots, \overline{t_m}) & = & \overline{F(t_1, \dots, t_m)} \\ R^{\mathcal{A}}(\overline{t_1}, \dots, \overline{t_n}) & \iff & \mathsf{T} \vdash R(t_1, \dots, t_n) \end{array}$$

Daß diese Definitionen “wohldefiniert” sind (d. h. nicht von der Wahl der Repräsentanten der jeweiligen Äquivalenzklassen abhängen), folgt wiederum aus den Gleichheitsaxiomen (Prinzip der Nichtunterscheidbarkeit, s. 4.4.2). Damit ist (für $K \neq \emptyset$ ) nun eine $\mathcal{L}$-Struktur

$$\mathcal{A} = \mathcal{A}_{\mathsf{T}} = (A, R^{\mathcal{A}}, \ldots F^{\mathcal{A}}, \ldots c^{\mathcal{A}}, \ldots) \text{ definiert,}$$

die **kanonische Struktur** der Theorie $\mathsf{T}$.

Eine Struktur bzw. ein Modell, dessen Elemente Interpretationen von konstanten Termen sind (wie die obige kanonische Struktur) nennt man auch eine **Termstruktur** bzw. ein **Termmodell**.

### 4.8.2 Satz über Termmodelle

*Für alle atomaren Sätze $\sigma$ gilt:*

$$(*) \quad \mathcal{A}_{\mathsf{T}} \models \sigma \iff \mathsf{T} \vdash \sigma.$$

*Beweis:* Zeige zunächst durch Induktion über den Termaufbau:

(1) $t^{\mathcal{A}} = \bar{t}$ für alle $t \in K$.

(*) beweist man dann durch Induktion über die Formellänge:

(2) Ist $\sigma = R(t_1, \ldots, t_n)$, so

$$\begin{array}{rcll} \mathcal{A}_{\mathsf{T}} \models \sigma & \iff & R^{\mathcal{A}}(t_1^{\mathcal{A}}, \ldots, t_n^{\mathcal{A}}) & \text{nach Def. von } \models \\ & \iff & R^{\mathcal{A}}(\overline{t_1}, \ldots, \overline{t_n}) & \text{nach(1)} \\ & \iff & \mathsf{T} \vdash R(t_1, \ldots, t_n) & \text{nach Def. von } R^{\mathcal{A}} \end{array}$$

(3) Ist $\sigma$ der Satz $s = t$, so

$$\begin{array}{rcll} \mathcal{A}_{\mathsf{T}} \models \sigma & \iff & s^{\mathcal{A}} = t^{\mathcal{A}} & \text{nach Def. von } \models \\ & \iff & \bar{s} = \bar{t} & \text{nach (1)} \\ & \iff & \mathsf{T} \vdash s = t & \text{nach Def. von } \bar{t} \quad \square \end{array}$$

Könnten wir die Aussage (*) für alle Sätze zeigen, so wäre die Struktur $\mathcal{A}_{\mathsf{T}}$ ein Modell von $\mathsf{T}$. Im allgemeinen gilt dies aber nicht, $\mathsf{T}$ kann zwar widerspruchsfrei sein, aber möglicherweise zu wenig Information enthalten: Um den Induktionsbeweis weiterführen zu können, müßte gelten:

(a) analog zu

$$\begin{aligned} \mathcal{A}_\mathsf{T} \models \neg\sigma &\iff \mathcal{A}_\mathsf{T} \not\models \sigma : \\ \mathsf{T} \vdash \neg\sigma &\iff \mathsf{T} \nvdash \sigma, \end{aligned}$$

(b) analog zu

$$\begin{aligned} \mathcal{A}_\mathsf{T} \models \exists x\varphi &\iff \textit{es ex. ein } t \in K \textit{ mit } \mathcal{A}_\mathsf{T} \models \varphi[\bar{t}/x] : \\ \mathsf{T} \vdash \exists x\varphi &\iff \textit{es ex. ein } t \in K \textit{ mit } \mathsf{T} \vdash \varphi(t). \end{aligned}$$

Die Eigenschaft (a) charakterisiert gerade die konsistenten und (syntaktisch) *vollständigen* Theorien, dagegen erfüllen *Henkin-Theorien* die Eigenschaft (b). Wir werden zeigen, dass man beide Eigenschaften durch eine geeignete Erweiterung von $\mathsf{T}$ erhalten kann.

## 4.9 Vervollständigung von Theorien

Wie im Falle der Aussagenlogik nennen wir eine Theorie $\mathsf{T}$ der Sprache $\mathcal{L}$ (syntaktisch) **vollständig** gdw

$$\textit{für jeden } \mathcal{L}\textit{-Satz } \sigma \textit{ gilt: } \mathsf{T} \vdash \sigma \textit{ oder } \mathsf{T} \vdash \neg\sigma.$$

Da $\mathsf{T}$ konsistent ist gdw nicht zugleich $\mathsf{T} \vdash \sigma$ und $\mathsf{T} \vdash \neg\sigma$ gelten, so ist $\mathsf{T}$ konsistent und (syntaktisch) vollständig gdw

$$\textit{für jeden } \mathcal{L}\textit{-Satz } \sigma \textit{ gilt: entweder } \mathsf{T} \vdash \sigma \textit{ oder } \mathsf{T} \vdash \neg\sigma.$$

Der Vollständigkeitsbegriff im semantischen Sinne (für jeden $\mathcal{L}$-Satz $\sigma$ gilt: $\mathsf{T} \models \sigma$ *oder* $\mathsf{T} \models \neg\sigma$) wird mit obigem übereinstimmen, sobald wir den Vollständigkeitssatz bewiesen haben. Solange soll *vollständig* hier im syntaktischen Sinne gemeint sein (und besser durch den äquivalenten Begriff *maximal konsistent* ersetzt werden). Daneben haben wir noch den Begriff des *vollständigen Axiomensystems*; dieser ist von den Begriffen der *vollständigen Theorie* streng zu unterscheiden. Daher beachte man:

Der Vollständigkeitssatz für die Prädikatenlogik besagt nicht,
daß jede Theorie vollständig ist!

Die Vollständigkeit eines Axiomensystems für die Logik bedeutet, daß es genügend Axiome und Regeln enthält, um alle wahren Sätze zu beweisen; bezüglich der Folgerungsbeziehung besagt der Vollständigkeitssatz, daß alle Sätze, die in allen Modellen einer Theorie $\mathsf{T}$ gelten, aus $\mathsf{T}$ mit den Axiomen und Regeln des Axiomensystems beweisbar sind. Daraus folgt keineswegs, daß alle Theorien vollständig sind, häufig sollen sie sogar unvollständig sein, um möglichst viele Modelle mit sehr unterschiedlichen Eigenschaften zuzulassen (wie die algebraische Theorien der Gruppen, Ringe oder Körper). Für den Beweis des Vollständigkeitssatzes der Prädikatenlogik benötigen wir allerdings den Begriff der vollständigen Theorie und die Möglichkeit, eine beliebige konsistente Theorie zu einer maximal konsistenten Theorie zu erweitern.

### 4.9.1 Satz von Lindenbaum

*$\mathsf{T}$ sei eine konsistente Theorie der abzählbaren Sprache $\mathfrak{L}$. Dann existiert eine maximal konsistente Erweiterung $\mathsf{T}_V$ von $\mathsf{T}$ in derselben Sprache $\mathfrak{L}$.*

*Beweis:* Da die Sprache $\mathfrak{L}$ abzählbar ist, existiert eine Aufzählung $(\sigma_n | n \in \mathbb{N})$ aller $\mathfrak{L}$-Sätze. Setze (rekursiv)

$$\mathsf{T}_0 = \mathsf{T} \quad \text{und} \quad \mathsf{T}_{n+1} = \begin{cases} \mathsf{T}_n, & \text{falls } \mathsf{T}_n \vdash \sigma_n, \\ \mathsf{T}_n[\neg\sigma_n] & \text{sonst.} \end{cases}$$

Schließlich sei $\quad \mathsf{T}_V := \bigcup_{n\in\mathbb{N}} \mathsf{T}_n.$

Da nach Konstruktion von $\mathsf{T}_{n+1}$ gilt:

$$\mathsf{T}_n \vdash \sigma_n \ \text{ oder } \ \mathsf{T}_{n+1} \vdash \neg\sigma_n, \quad \text{so auch}$$
$$\mathsf{T}_V \vdash \sigma_n \ \text{ oder } \ \mathsf{T}_V \vdash \neg\sigma_n,$$

und damit ist $\mathsf{T}_V$ eine vollständige Theorie.

Zum Beweis der Konsistenz von $\mathsf{T}_V$ zeigt man durch Induktion, daß alle $\mathsf{T}_n$ konsistent sind: Ist $\mathsf{T}_n$ widerspruchsfrei, so auch $\mathsf{T}_{n+1}$, wobei man im "sonst"-Fall 4.7.4 benutzt). Die Widerspruchsfreiheit von $\mathsf{T}_V$ folgt dann aus dem Endlichkeitssatz für die Beweisbarkeit. □

**Bemerkungen**

- Der Beweis des Satzes von LINDENBAUM zeigt, daß die Existenz einer Vervollständigung von der Wahl der Aufzählung der Sätze der Sprache abhängt;

i. a. hat eine (konsistente aber selbst nicht vollständige) Theorie sehr viele Vervollständigungen!

- $\mathsf{T}$ ist konsistent und vollständig gdw $\mathsf{C}(\mathsf{T}) := \{\sigma \mid \mathsf{T} \vdash \sigma\}$ (die Menge aller Folgerungen aus $\mathsf{T}$) eine maximale konsistente Theorie ist.

- Um den Satz von LINDENBAUM für beliebige Sprachen zu beweisen, muß man also zeigen, daß sich jede konsistente Theorie zu einer maximalen konsistenten Theorie erweitern läßt. Dazu benötigt man (zumindest im Fall überabzählbarer Sprachen) das Auswahlaxiom, etwa in Form des ZORNschen Lemmas (s. 7.2).

## 4.10 Henkin-Theorien

Wir behandeln jetzt Theorien, in denen der Existenzquantor in Sätzen eliminiert werden kann, und zwar durch eine Konstante als **Beispiel**:

### Definition einer Henkin-Theorie

Eine Theorie $\mathsf{T}$ der Sprache $\mathcal{L}$ heißt **Henkin-Theorie** gdw für jeden $\mathcal{L}$-Satz der Form $\sigma = \exists x\varphi$ gilt: Es gibt einen konstanten Term $t$ der Sprache $\mathcal{L}$ mit

$$\mathsf{T} \vdash \exists x\varphi \rightarrow \varphi(t) \text{ und somit}$$
$$\mathsf{T} \vdash \exists x\varphi \leftrightarrow \varphi(t).$$

$\mathcal{A}$ sei $\mathcal{L}$-Struktur mit Individuenbereich $A$. Bereits früher haben wir die Sprache $\mathcal{L}_A$ eingeführt, die zusätzliche Namen $\underline{a}$ für jedes Element $a \in A$ enthält. In der Struktur $\mathcal{A}_A$ wird $\underline{a}$ durch $a$ selbst interpretiert, und mit

$$\mathsf{D}_{\mathcal{L}}(\mathcal{A}) = \mathsf{Th}(\mathcal{A}_A) = \{\sigma \mid \mathcal{A}_A \models \sigma, \sigma\, \mathcal{L}_A\text{-}Satz\}$$

haben wir das **elementare Diagramm** von $\mathcal{A}$ bezeichnet.

### Beispiele

(i) $\mathsf{Th}(\mathcal{A})$ und $\mathsf{D}_{\mathcal{L}}(\mathcal{A})$ sind stets konsistente und vollständige Theorien, letztere ist offenbar eine HENKIN-Theorie. Als Folgerung aus dem Vollständigkeitssatz wird sich ergeben, daß umgekehrt jede konsistente und vollständige Theorie von der Form $\mathsf{Th}(\mathcal{A})$ ist für ein $\mathcal{A}$.

(ii) Konkrete Beispiele sind

- die Theorie der *dichten linearen Ordnung ohne Endpunkte:* konsistent und vollständig (s. 9.2), aber (in der Sprache allein mit $<$-Zeichen) keine HENKIN-Theorie,
- die *Gruppentheorie* ist konsistent, aber unvollständig: Das Kommutativgesetz $\forall x \forall y (x \circ y = y \circ x)$ ist weder beweisbar noch widerlegbar, aber auch die Theorie der ABELschen Gruppen ist offenbar noch unvollständig.
- Die PEANO-Arithmetik $\mathsf{PA}$ ist unvollständig (GÖDELscher Unvollständigkeitssatz).
- Dagegen ist die Theorie des Standardmodells von $\mathsf{PA}$, $\mathsf{Th}(\mathbb{N},',0)$, eine vollständige HENKIN-Theorie.

Eine HENKIN-Theorie muß zu jedem Existenzsatz einen konstanten Term als Beispiel besitzen, i. a. besitzen natürlich vorgegebene Theorien nicht genügend derartige Beispiele. Man kann jedoch durch Hinzunahme genügend vieler Konstanten eine Theorie zu einer HENKIN-Theorie erweitern:

### Henkin-Erweiterungen

Es sei $\mathsf{T}$ eine Theorie der Sprache $\mathcal{L}$. Wähle für jeden $\mathcal{L}$-Satz der Form $\sigma = \exists x \psi$ eine neue (d. h. in der Sprache $\mathcal{L}$ bisher nicht vorkommende) Konstante $c_\sigma$ und bilde die erweiterte Sprache

$$\begin{aligned} \mathcal{L}^* &:= \mathcal{L} \cup \{c_\sigma \mid \sigma\, \mathcal{L}\textit{-Satz der Form } \exists x \psi\}, \\ \mathsf{T}^* &:= \mathsf{T}[\Gamma] \quad \text{mit den HENKIN-Axiomen in der Sprache } \mathcal{L}^*: \\ \Gamma &:= \{\exists x \psi(x) \to \psi(c_\sigma) \mid \sigma\, \mathcal{L}\textit{-Satz der Form } \exists x \psi\}. \end{aligned}$$

$\mathsf{T}^*$ braucht aber noch keine HENKIN-Theorie zu sein, da man mit Hilfe der neuen Konstanten möglicherweise neue Existenzsätze finden kann, für die keine Beispiele in der Sprache $\mathcal{L}^*$ vorhanden sind. Man muß daher diesen Prozeß unendlich oft iterieren: Setze

$$\begin{aligned} \mathcal{L}_0 = \mathcal{L}\,, &\quad \mathsf{T}_0 = \mathsf{T} \quad \text{Theorie der Sprache } \mathcal{L} \\ \mathcal{L}_{n+1} = \mathcal{L}_n^*, &\quad \mathsf{T}_{n+1} = \mathsf{T}_n^* \quad \text{Theorie der Sprache } \mathcal{L}_{n+1}. \end{aligned}$$

Schließlich sei

$\mathcal{L}_H = \bigcup_{n\in\mathbb{N}} \mathcal{L}_n, \quad \mathsf{T}_H = \bigcup_{n\in\mathbb{N}} \mathsf{T}_n$ die **Henkin-Erweiterung** der Theorie $\mathsf{T}$.

**Satz über Henkin-Erweiterungen**

(i) $\mathsf{T}_H$ *ist eine* HENKIN-*Theorie,*

(ii) $\mathsf{T}_H$ *ist eine konservative Erweiterung von* $\mathsf{T}$,

(iii) *ist die Sprache* $\mathcal{L}$ *abzählbar, so auch die Sprache* $\mathcal{L}_H$.

*Beweis:* (i) und (iii) sind klar (Abzählbarkeit einer Sprache bedeutet, daß die Menge der Symbole der Sprache abzählbar ist). Zum Beweis von (ii) sei $\mathsf{T}'$ die Theorie der Sprache $\mathcal{L}_H$ mit denselben nicht-logischen Axiomen wie $\mathsf{T}$. Dann ist $\mathsf{T}'$ eine rein sprachliche Erweiterung von $\mathsf{T}$ und somit konservative Erweiterung von $\mathsf{T}$. Wir zeigen nun die Behauptung, indem wir nachweisen:

(+) *ist* $\varphi$ *Formel der Sprache* $\mathcal{L}$ *und gilt* $\mathsf{T}_H \vdash \varphi$*, so auch* $\mathsf{T}' \vdash \varphi$.

Sei also $\varphi$ $\mathcal{L}$-Formel mit $\mathsf{T}_H \vdash \varphi$. Dann gibt es endlich viele Sätze der Form $\sigma_i = \exists x \psi_i(x) \to \psi_i(c_i)$ mit $c_i = c_{\exists x \psi_i} (i = 1, \ldots, n)$, so daß

$$\mathsf{T}' \vdash \sigma_1 \wedge \ldots \wedge \sigma_n \to \varphi.$$

Wir zeigen $\mathsf{T}' \vdash \varphi$ durch Induktion nach $n$:

Jeder Satz $\sigma_i$ ist ein Satz der Sprache $\mathcal{L}_{m_i}$ für ein $m_i$, wobei wir annehmen können, daß $m_1 \geq m_2, \ldots, m_n$. Das bedeutet dann insbesondere, daß die Konstante $c_1$ in den Sätzen $\sigma_2, \ldots, \sigma_n$ und damit auch in den Formeln $\psi_1, \ldots, \psi_n, \varphi$ nicht vorkommt. Also existiert eine Variable $y$, die nicht in diesen Formeln vorkommt, mit (s. Beweis von 4.7.2)

| | |
|---|---|
| $\mathsf{T}' \vdash (\exists x \psi_1(x) \to \psi_1(y)) \wedge \sigma_2 \wedge \ldots \wedge \sigma_n \to \varphi.$ | Dann: |
| $\mathsf{T}' \vdash (\exists x \psi_1(x) \to \psi_1(y)) \to (\sigma_2 \wedge \ldots \wedge \sigma_n \to \varphi)$ | und mit $\exists$-Einführung |
| $\mathsf{T}' \vdash \exists y (\exists x \psi_1(x) \to \psi_1(y)) \to (\sigma_2 \wedge \ldots \wedge \sigma_n \to \varphi).$ | Nun gilt aber |
| $\vdash \exists y (\exists x \psi_1(x) \to \psi_1(y)),$ | und somit haben wir |
| $\mathsf{T}' \vdash \sigma_2 \wedge \ldots \wedge \sigma_n \to \varphi.$ | |

Ähnlich (bzw. nach Ind.vor.) kann man die übrigen Prämissen eliminieren und erhält schließlich $\mathsf{T}' \vdash \varphi$. $\square$

## 4.11 Der Gödelsche Vollständigkeitssatz

Dieser Satz wurde zuerst von Kurt GÖDEL für abzählbare Sprachen bewiesen:

**Gödel, Kurt** *Die Vollständigkeit der Axiome des logischen Funktionenkalküls.* Monatshefte Math. Phys. 37 (**1930**), 349-360.

und für beliebige Sprachen 1936 von I. MALCEV. Unsere Beweismethode geht zurück auf

**Henkin, Leon** *The completeness of the first-order functional calculus.* Journal of Symbolic Logic 14 (**1949**), 159 - 166, s.a.:

**Henkin, Leon** *The discovery of my completeness proof.* Bull. Journal of Symbolic Logic 2,2 (**1996**), 127 - 158,

**Leblanc-Roeper-Than-Weaver** *Henkin´s completeness proof: Forty years later.* Notre Dame Journal of Formal Logic 32 (**1991**), 212 - 232.

Die HENKINsche Methode hat den Vorzug, daß man sie vielfältig abwandeln kann, um - in Abhängigkeit von speziellen Eigenschaften einer Theorie $\mathsf{T}$ - Modelle von $\mathsf{T}$ mit besonderen Eigenschaften zu erhalten. Wie schon erwähnt, läßt sich der Vollständigkeitssatz in zwei äquivalenten Fassungen angeben:

### 4.11.1 Vollständigkeit der Prädikatenlogik / Modellexistenz-Satz

(1) $\mathsf{T} \models \sigma \Longrightarrow \mathsf{T} \vdash \sigma$,

(2) $\mathsf{T}$ *konsistent* $\Longrightarrow$ $\mathsf{T}$ *erfüllbar.*

*Beweis:* Wie im Falle der Aussagenlogik (siehe auch 4.7.4) folgt (1) aus (2) (die Umkehrung ist noch einfacher zu zeigen), so daß wir nur (2) zu beweisen brauchen.

Sei also $\mathsf{T}$ eine konsistente Theorie der Sprache $\mathcal{L}$. Wir erweitern $\mathsf{T}$ zunächst zu einer HENKIN-Theorie $\mathsf{T}_H$ in der Sprache $\mathcal{L}' := \mathcal{L}_H$ und bilden dann eine Vervollständigung dieser Theorie:

$$\mathsf{T} \subseteq \mathsf{T}_H \subseteq (\mathsf{T}_H)_V = \mathsf{T}'$$

Da beim zweiten Schritt die Sprache nicht mehr erweitert wird, ist $\mathsf{T}'$ nicht nur vollständig und konsistent, sondern auch eine HENKIN-Theorie geblieben. Nun wählen wir das Termmodell $\mathcal{A}'$ von $\mathsf{T}'$ und zeigen, daß für alle Sätze $\sigma$ der Sprache $\mathcal{L}'$ gilt:

$$(*) \quad \mathcal{A}' \models \sigma \iff \mathsf{T}' \vdash \sigma.$$

Insbesondere ist dann $\mathcal{A}'$ ein Modell von $\mathsf{T}$ und damit $\mathcal{A} := \mathcal{A}' \restriction \mathcal{L}$, die Einschränkung auf die Sprache $\mathcal{L}$, ein Modell von $\mathsf{T}$. Die Behauptung (*) zeigen wir durch Induktion über den Aufbau von $\sigma$:

1. Fall: $\sigma$ ist atomar. Dieser Anfangsfall ist gerade Inhalt des Satzes 4.8.2 über Termmodelle.

2. Fall: $\sigma = \neg\psi$. Dann gilt nach Definition von $\models$:

$$\begin{aligned} \mathcal{A}' \models \neg\psi &\iff \mathcal{A}' \not\models \psi \quad \text{und entsprechend} \\ \mathsf{T}' \vdash \neg\psi &\iff \mathsf{T}' \not\vdash \psi, \end{aligned}$$

wegen der Vollständigkeit und Konsistenz von $\mathsf{T}'$. Somit folgt die Behauptung aus der Induktionsvoraussetzung.

3. Fall: $\sigma = \psi \vee \theta$. Dann gilt wieder nach der Definition von $\models$:

$$\begin{aligned} \mathcal{A}' \models \psi \vee \theta &\iff \mathcal{A}' \models \psi \text{ oder } \mathcal{A}' \models \theta \quad \text{und entsprechend} \\ (**) \quad \mathsf{T}' \vdash \psi \vee \theta &\iff \mathsf{T}' \vdash \psi \text{ oder } \mathsf{T}' \vdash \theta, \end{aligned}$$

denn in (**) gilt "$\Rightarrow$" mittels aussagenlogischer Abschwächung. Für den Beweis von "$\Leftarrow$" nehmen wir an, daß

$$\begin{array}{ccc} \mathsf{T}' \vdash \psi & \text{oder } \mathsf{T}' \vdash \theta & \textit{nicht} \text{ gilt, also} \\ \mathsf{T}' \not\vdash \psi & \text{und } \mathsf{T}' \not\vdash \theta & \text{und somit} \\ \mathsf{T}' \vdash \neg\psi & \text{und } \mathsf{T}' \vdash \neg\theta & \end{array}$$

wegen der Vollständigkeit von $\mathsf{T}'$. Dann gilt aber als tautologische Folgerung: $\mathsf{T}' \vdash \neg(\psi \vee \theta)$, und das heißt $\mathsf{T}' \not\vdash \psi \vee \theta$ wegen der Konsistenz von $\mathsf{T}'$.

4. Fall: $\sigma = \exists x\psi$. Dann gilt wieder nach der Definition von $\models$:

$$\mathcal{A}' \models \exists x\psi \iff \text{es ex. ein } a \in |\mathcal{A}'| \text{ mit } \mathcal{A}' \models \psi[a].$$

Da $\mathcal{A}'$ ein Termmodell ist, so ist jedes $a \in |\mathcal{A}'|$ von der Form $a = \bar{t} = t^{\mathcal{A}'}$ für einen konstanten Term $t$, also:

$$\mathcal{A}' \models \exists x \psi \iff \text{es ex. ein } t \in K \text{ mit } \mathcal{A}' \models \psi[t/x].$$

Entsprechend gilt, da $\mathsf{T}'$ HENKIN-Theorie ist:

$$\mathsf{T}' \vdash \exists x \psi \iff \text{es ex. ein } t \in K \text{ mit } \mathsf{T}' \vdash \psi(t).$$

Somit folgt die Behauptung auch in diesem Fall aus der Induktionsvoraussetzung. □

Im Falle einer abzählbaren Sprache läßt sich der Satz von LINDENBAUM und damit auch der Vollständigkeitssatz ohne Voraussetzung des Auswahlaxioms beweisen. Auf diesen Fall bezieht sich als Korollar auch der

### 4.11.2 Satz von Löwenheim

*Ist $\mathsf{T}$ eine konsistente Theorie in einer abzählbaren Sprache,*
*so besitzt $\mathsf{T}$ ein abzählbares Modell.*

Das Ergebnis von LÖWENHEIM findet sich in der Form

*Hat eine Theorie $\mathsf{T}$ in einer abzählbaren Sprache $\mathcal{L}$ ein Modell,*
*so auch ein abzählbares Modell.*

bereits in der Arbeit

**Löwenheim, Leopold** *Über Möglichkeiten im Relativkalkül.* Math. Ann. 76, (**1915**), 447 - 470.

Verallgemeinerungen werden wir später mit modelltheoretischen Methoden nach SKOLEM-TARSKI behandeln (10.4 und 10.5).

Aus dem Vollständigkeitssatz folgt auch das

### Korollar

(i) *Sind $\mathsf{T}$ und $\mathsf{T}'$ Theorien in derselben Sprache $\mathcal{L}$, so gilt:*

$$\mathsf{T} \subseteq \mathsf{T}' \iff Mod(\mathsf{T}) \supseteq Mod(\mathsf{T}'),$$
$$\mathsf{T} = \mathsf{T}' \iff Mod(\mathsf{T}) = Mod(\mathsf{T}').$$

(ii) *Eine Theorie $\mathsf{T}$ ist konsistent und vollständig gdw es eine Struktur $\mathcal{A}$ gibt mit $\mathsf{T} = \mathsf{Th}(\mathcal{A})$.*

Dabei meinen wir hier mit $\mathsf{T} = \mathsf{T}'$ : $\mathsf{T}$ äquivalent zu $\mathsf{T}'$, d. h. $\mathsf{C}(\mathsf{T}) = \mathsf{C}(\mathsf{T}')$. □

Als weitere wichtige Folgerung aus dem Vollständigkeitssatz (und dem Endlichkeitssatz für die Beweisbarkeit) erhalten wir erneut den

### 4.11.3 Kompaktheitssatz

$\Sigma$ *sei eine Menge von Sätzen,* $\sigma$ *ein Satz. Dann gilt:*

(1) $\mathsf{T} \models \sigma \Rightarrow$ *es existiert ein endliches* $\mathsf{T}_0 \subseteq \mathsf{T}$ mit $\mathsf{T}_0 \models \sigma$,

(2) *hat jedes endliche* $\Sigma_0 \subseteq \Sigma$ *ein Modell, so hat auch* $\Sigma$ *ein Modell.*

(Natürlich gelten auch die jeweiligen Umkehrungen.) □

**Bemerkung**

Folgende Ergebnisse sind äquivalent:

(i) Der Vollständigkeitssatz (für beliebige Sprachen),

(ii) der Satz von LINDENBAUM (für beliebige Sprachen),

(iii) der Kompaktheitssatz (für beliebige Sprachen),

(iv) das BOOLEsche Primidealtheorem:

*Jedes echte Ideal einer* BOOLE*schen Algebra läßt sich zu einem Primideal erweitern (bzw. jeder echte Filter zu einem Ultrafilter).*

Alle diese Aussagen folgen aus dem Auswahlaxiom, sind aber schwächer als dieses.

# Teil II

# Mengenlehre

# Kapitel 5

# Grundlagen der Mengenlehre

Quod enim dicimus aliquid esse infinitum,
non aliquid in re significamus,
sed impotentiam in animo nostro;
tanquam si diceremus, nescire nos, an et ubi
terminetur.
*Th. Hobbes, De cive XV 14*

...l´homme qui n´est produit que pour l´infinité.
*Blaise Pascal: Préface sur le Traité du vide*

## 5.1 Ordinalzahlen

In seinen Untersuchungen über die Konvergenz von Fourierreihen führte CANTOR den Begriff der *Ableitung* $A'$ einer Menge $A$ von reellen Zahlen ein: $A'$ besteht aus den Elementen von $A$, welche Häufungspunkte von Elementen von $A$ sind. Man kann nun nach der Menge der Häufungspunkte dieser neuen Menge fragen und damit eine Menge $A''$ bilden, die Menge der Häufungspunkte der Häufungspunkte von $A$. Iteriert man diese Operation, so erhält man eine eine unendlich-absteigende Folge

$$A \supseteq A' \supseteq A'' \supseteq A''' \ldots$$

deren "Limes" $A^\infty$, d. h. in diesem Fall der Durchschnitt, möglicherweise wieder isolierte Punkte enthält, so daß man wieder die Menge ihrer Häufungspunkte bilden kann. Auf diese Weise fortgesetzt, führt der Zählprozeß über die natürlichen Zahlen hinaus ins Transfinite:

$0, 1, 2, 3, \dots \infty, \infty + 1, \infty + 2, \infty + 3, \dots$
$\infty + \infty, \infty + \infty + 1, \infty + \infty + 2, \dots \infty + \infty + \infty \dots$

Einen derartigen Zählprozeß erhalten wir auch, wenn wir die natürlichen Zahlen umordnen und neu aufzählen, etwa erst die Potenzen von 2, dann Potenzen von 3, dann die Potenzen von 5 ...:

$$1, 2, 4, 8, \dots 3, 9, 27, \dots 5, 25, 125, \dots$$

aber wo bleiben etwa die Zahlen 0, 6, 10, . . . ? Offenbar gibt es verschiedene Möglichkeiten, ins Unendliche aufzuzählen, wie unterscheiden sich diese Möglichkeiten? Führen unterschiedliche Aufzählungen vielleicht zu Widersprüchen? Lassen sich überhaupt alle Mengen in irgendeiner Weise aufzählen? Tatsächlich ergeben sich Widersprüche, wenn man allzu naiv versucht, Eigenschaften von endlichen auf unendliche Mengen zu übertragen; trotzdem kann man aber die Theorie der Wohlordnungen und der Ordinalzahlen einheitlich für endliche und unendliche Mengen begründen. Hier setzen wir zunächst einen intuitiven Mengenbegriff voraus (eine formale Begründung werden wir später nachliefern) und erinnern an die in der Mathematik üblichen Ordnungsbegriffe:

### 5.1.1 Ordnungen

1. Eine (irreflexive) **teilweise Ordnung** auf einer Menge $a$ ist eine 2-stellige Relation $<$ auf $a$, für die gilt:

   (a) $\forall x \in a \quad x \not< x$ **irreflexiv**,
   (b) $\forall x, y, z \in a \, (x < y \wedge y < z \rightarrow x < z)$ **transitiv**.

2. Eine (irreflexive) **lineare Ordnung** auf $a$ ist eine teilweise Ordnung auf $a$ mit
   (c) $\forall x, y \in a \, (x < y \vee x = y \vee y < x)$ **vergleichbar**.

Eine teilweise Ordnung nennt man manchmal auch eine **partielle** (oder **Halb-**) Ordnung, eine lineare Ordnung manchmal auch einfach **Ordnung**. Die zugehörigen reflexiven Ordnungen erhält man durch die Festlegung

$$a \leq b \leftrightarrow a < b \vee a = b,$$

und umgekehrt kann man die obigen Begriffe auch erst axiomatisch für die reflexiven Varianten charakterisieren und die zugehörigen irreflexiven durch

$$a < b \leftrightarrow a \leq b \wedge a \neq b$$

definieren.

Die gewöhnlichen Ordnungen auf den natürlichen, den ganzen, den rationalen und den reellen Zahlen sind offenbar lineare Ordnungen; die Relation

$$f < g :\leftrightarrow \forall x \in \mathbb{R} \quad f(x) < g(x)$$

auf den reellen Funktionen ist dagegen nur eine teilweise Ordnung.

Die natürlichen Zahlen lassen sich der Größe nach aufzählen, aber für die anderen Zahlbereiche ist dies nicht möglich; selbst wenn man noch $-\infty$ als "kleinste Zahl" hinzunimmt, gibt es keine nächstgrößere (und bei dichten Ordnungen wie den rationalen Zahlen gibt es zu überhaupt keiner Zahl eine nächstgrößere). Um diese Bereiche in der Form $\{a_0, a_1, \ldots a_n, a_{n+1} \ldots\}$ aufzuzählen, müssen wir sie neu ordnen, so daß man mit

(i) einem kleinsten Element beginnen kann,
und wenn man in der Aufzählung zu einem Element gekommen ist,

(ii) weiß, mit welchem Element man fortfahren kann, und schließlich

(iii) auch den Aufzählungsprozeß fortsetzen kann, wenn man bereits eine unendliche Teilfolge von Elementen aufgezählt hat (aber noch nicht alles aufgezählt ist).

Diese Anforderungen kann man präzisieren und zugleich vereinheitlichen, indem man verlangt, daß jede nicht-leere Teilmenge (nämlich der Rest der noch nicht aufgezählten Elemente) ein kleinstes Element enthält (welches als "nächstes" aufzuzählen ist):

3. Eine **Wohlordnung** auf einer Menge $a$ ist eine lineare Ordnung auf $a$, welche zusätzlich die **Minimalitätsbedingung**

   (Min) $\forall z(\emptyset \neq z \subseteq a \rightarrow \exists x \in z\, \forall y \in z\, y \not< x)$

   erfüllt, welche wegen der Vergleichbarkeit (c) äquivalent ist zur **Existenz eines kleinsten Elementes**:

   (Kl) $\forall z(\emptyset \neq z \subseteq a \rightarrow \exists x \in z\, \forall y \in z\, x \leq y)$,
   wobei wie üblich $x \leq y :\leftrightarrow x < y \vee x = y$.

**Beispiele**

Jede lineare Wohlordnung auf einer endlichen Menge ist eine Wohlordnung, ebenso die gewöhnliche Ordnung auf den natürlichen Zahlen. Dagegen sind die ganzen Zahlen erst wohlgeordnet, wenn wir sie (etwa) in folgende Ordnung bringen:

$$0,1,2,3,4,\dots,-1,-2,-3,-4,\dots \text{ oder kürzer:}$$
$$0,1,-1,2,-2,3,-3,\dots$$

Im ersten Fall werden wir von einer Ordnung vom Typ $\omega+\omega$ , sprechen, im zweiten Fall vom Typ $\omega$ (und viele andere Möglichkeiten noch kennenlernen). Dabei ist das frühere Symbol $\infty$ durch $\omega$ ersetzt worden, zugleich das einfachste (unendliche) Beispiel der jetzt einzuführenden Ordinalzahlen. Diese wurden von CANTOR als Repräsentanten (isomorpher) Wohlordnungen eingeführt; heute definiert man sie nach VON NEUMANN als Mengen, die durch die $\in$-Beziehung wohlgeordnet sind und außerdem transitiv sind:

**Definition**

$$\begin{array}{rcl} trans(a) :\leftrightarrow & \forall x \in a \forall y \in x\, y \in a & \qquad \textbf{transitiv} \\ \leftrightarrow & \forall x \in a\, x \subseteq a & \end{array}$$

*Achtung: $trans(a)$* bedeutet *nicht*, daß die $\in$-Beziehung auf $a$ transitiv ist!

**Beispiele**

$\{\{\emptyset\}\}$ ist nicht transitiv, aber $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots$ sind transitiv und werden als Definition der natürlichen Zahlen $0,1,2,\dots$ gewählt:

$$\begin{array}{ll} 0 := \emptyset, & \\ a' := a+1 := a \cup \{a\} & \qquad \textbf{Nachfolger} \text{ von } a. \end{array}$$

Beginnt man also mit $0 := \emptyset$ und wendet wiederholt die Nachfolgeroperation an, so erhält man

$1 := \{\emptyset\},\ 2 := \{\emptyset, \{\emptyset\}\} = \{0,1\},\ 3 := \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = \{0,1,2\}\dots$
allgemein: $\quad n' = n \cup \{n\} = \{0,1,2,\dots,n\}$

für jede natürliche Zahl $n$ (was wir aber erst noch definieren müssen). Diese Definition hat den Vorteil hat, daß

- die $<$-Beziehung auf den natürlichen Zahlen mengentheoretisch besonders einfach ist, nämlich die $\in$-Relation, und
- jede natürliche Zahl $n$ genau $n$-viele Elemente enthält (nämlich die kleineren, beginnend mit 0 und endend mit $n-1$).

### 5.1.2 Definition der Ordinalzahlen

$$\begin{array}{rll}
con(a) :\leftrightarrow & \forall x,y \in a(x \in y \vee x = y \vee y \in x) & \textbf{connex}, \\
fund(a) :\leftrightarrow & \forall x \subseteq a(x \neq \emptyset \to \exists y \in x\, y \cap x = \emptyset) & \textbf{fundiert}, \\
Ord(a) :\leftrightarrow & trans(a) \wedge con(a) \wedge fund(a) & \textbf{Ordinalzahl}, \\
\in_a := & \{x,y \mid x,y \in a \wedge x \in y\} & \text{Elementbeziehung auf } a.
\end{array}$$

**Vereinbarung zur Fundierung**

Wir werden der Einfachheit halber voraussetzen, daß *alle Mengen fundiert sind*. Dann gilt also insbesondere:

$$b \neq \emptyset \to \exists y \in b\; y \cap b = \emptyset, \text{ d. h.}$$
$$b \neq \emptyset \to \exists y \in b\, \forall z \in b\; z \notin y,$$

was besagt, daß jede nicht-leere Menge $b$ ein (bezüglich der $\in$-Relation) minimales Element besitzt (vgl. die Bedingung (Min) in der Definition einer Wohlordnung). Damit werden einige ungewöhnliche Mengen ausgeschlossen, insbesondere Mengen $a$, die sich selbst als Element enthalten oder mit anderen Mengen einen endlichen $\in$-Zyklus bilden. Es gilt somit:

$$(\mathsf{F}^*) \quad (a \notin a),\ \neg(b \in c \wedge c \in b),\ \neg(d \in b \wedge b \in c \wedge c \in d), \ldots$$

Damit läßt sich die Definition der Ordinalzahlen vereinfachen (siehe (i) in folgendem Satz). Später weren wir ohnehin das Fundierungsaxiom (in der Form: *alle Mengen sind fundiert*) voraussetzen; alle folgenden Aussagen über Ordinalzahlen lassen sich jedoch ohne das Fundierungsaxiom beweisen, wenn man die ursprüngliche Definition zugrunde legt.

### 5.1.3 Satz: Charakterisierung von Ordinalzahlen

(i) $Ord(a) \leftrightarrow trans(a) \wedge con(a)$.

(ii) *Ordinalzahlen sind transitive Mengen, die durch die $\in$-Relation wohlgeordnet sind:*
$Ord(a) :\leftrightarrow trans(a) \wedge \in_a$ ist Wohlordnung auf $a$.

(iii) *Elemente von Ordinalzahlen sind wieder Ordinalzahlen:*
$Ord(a) \wedge b \in a \rightarrow Ord(b)$.

*Beweis:* (i) gilt wegen unserer Vereinbarung. Für (ii) ist nur zu zeigen:

$$Ord(a) \rightarrow trans(\in_a).$$

Sei also $Ord(a)$, sowie $b, c, d \in a$ mit $b \in c \wedge c \in d$. Beh.: $b \in d$.

| | | |
|---|---|---|
| Es gilt: | $b \in d \vee d = b \vee d \in b$ | wegen $con(a)$. |
| Falls | $d = b$, so hätten wir $b \in c \wedge c \in b$ | im Widerspruch zu ($\mathsf{F}$*) |
| falls | $d \in b$, so $d \in b \wedge b \in c \wedge c \in d$ | im Widerspruch zu ($\mathsf{F}$*). |

Somit bleibt nur die Möglichkeit $b \in d$.

Um (iii) zu zeigen, sei $Ord(a) \wedge b \in a$. Dann ist wegen $trans(a) : b \subseteq a$ und mit $\in_a$ auch $\in_b$ eine Wohlordnung. Somit brauchen wir wegen (ii) nur noch zu zeigen, daß auch $b$ transitiv ist:
Sei $x \in y \in b$. Beh.: $x \in b$. Dazu argumentieren wir wie oben: Es sind $x, b \in a$ (ersteres wegen $trans(a)$), also sind beide wegen $con(a)$ miteinander vergleichbar: $x \in b \vee x = b \vee b \in x$, wobei die letzten beiden Möglichkeiten zu einem Widerspruch zur Fundierung ($\mathsf{F}$*) führen. □

### 5.1.4 Die Ordnung der Ordinalzahlen

Ordinalzahlen werden üblicherweise mit kleinen griechischen Buchstaben bezeichnet:

| | |
|---|---|
| $\alpha, \beta, \gamma, \ldots, \xi, \eta, \zeta, \ldots$ | stehen für Ordinalzahlen, ebenso die Quantoren |
| $\forall \xi \ \ldots \exists \zeta, \ \ldots$ | für $\forall x(Ord(x) \rightarrow \ldots), \exists y(Ord(y) \wedge \ldots), \ldots$ |

Ferner schreiben wir

| | |
|---|---|
| $\alpha < \beta$ | für $\alpha \in \beta$, |
| $\alpha \leq \beta$ | für $\alpha \in \beta \vee \alpha = \beta$. |

Wir werden gleich zeigen, daß die Ordinalzahlen durch die $\in$-Relation nicht nur geordnet, sondern sogar wohlgeordnet ist, so daß diese Bezeichnungsweise gerechtfertigt ist. Der obige Satz 5.1.3 besagt also im Teil (iii) : $\alpha = \{\xi \mid \xi < \alpha\}$.

**Satz**

$$trans(b) \wedge b \subseteq \alpha \rightarrow Ord(b) \wedge (b \in \alpha \vee b = \alpha), \textit{ insbesondere}$$
$$\beta \leq \alpha \leftrightarrow \beta \in \alpha \vee \beta = \alpha \leftrightarrow \beta \subseteq \alpha.$$

*Beweis:* Sei $trans(b) \wedge b \subseteq \alpha$. Dann ist auch $\in_b$ eine Wohlordnung, also $Ord(b)$ nach Satz 5.1.3 (ii).

Falls $b \neq \alpha$, so $b \subset \alpha$, d. h. $\alpha - b \neq \emptyset$, und wir können wegen der Fundiertheit ein minimales $\gamma \in \alpha - b$ wählen, von dem wir zeigen werden, daß $b = \gamma$ und damit $b \in \alpha$ wie erwünscht:

Sei also $\gamma \in \alpha - b$ $\in$-minimal, so daß insbesondere $\forall x \in \gamma\, x \in b$, d. h. $\gamma \subseteq b$. Es gilt dann aber auch $b \subseteq \gamma$:

Sei $x \in b$. Dann $x \in \alpha$ (nach Voraussetzung) und $x \in \gamma \vee x = \gamma \vee \gamma \in x$ wegen $con(\alpha)$. Aber die letzten beiden Fälle können nicht eintreten: $x = \gamma \rightarrow \gamma \in b$ und $\gamma \in x \rightarrow \gamma \in b$ (wegen $trans(b)$), es ist aber nach Wahl von $\gamma$: $\gamma \in \alpha - b$. □

Somit haben wir mengentheoretisch nicht nur eine einfache $<$-Beziehung auf den Ordinalzahlen (nämlich die $\in$-Beziehung), sondern auch eine einfache $\leq$-Beziehung (nämlich die $\subseteq$-Beziehung). Insbesondere gilt somit:

$$\alpha \cup \beta = \max\{\alpha, \beta\} \quad \textit{und} \quad \alpha \cap \beta = \min\{\alpha, \beta\}.$$

Es ist nun leicht zu sehen, daß 0 die kleinste Ordinalzahl ist und daß mit einer Ordinalzahl $\alpha$ auch $\alpha'$ wieder eine Ordinalzahl ist (und zwar die nächst größere), so daß alle natürlichen Zahlen insbesondere Ordinalzahlen sind. Es gibt aber tatsächlich unmäßig viele Ordinalzahlen:

**Satz**

(i) *Die $\in$-Beziehung ist eine Wohlordnung auf allen Ordinalzahlen.*

(ii) *Es gibt keine Menge aller Ordinalzahlen.*
(**Antinomie von Burali-Forti**)

*Beweis* von (i):

| | |
|---|---|
| $\alpha \notin \alpha$ | nach dem Fundierungsaxiom (oder wegen $fund(\alpha)$) |
| $\alpha \in \beta \in \gamma \rightarrow \alpha \in \gamma$ | wegen $trans(\gamma)$ |
| $\alpha \in \beta \vee \alpha = \beta \vee \beta \in \alpha$ | kann man wie folgt zeigen: |

Sei $\delta := \alpha \cap \beta$ das Minimum der beiden Ordinalzahlen. Dann ist $trans(\delta)$ und $\delta \subseteq \alpha, \beta$, also nach nach dem vorangegegangen Satz: $\delta = \alpha \vee \delta \in \alpha$ und ebenso $\delta = \beta \vee \delta \in \beta$, aber im Fall $\delta \in \alpha \wedge \delta \in \beta$ erhielten wir den Widerspruch $\delta \in \delta$!

Somit ist die $\in$-Beziehung auf den Ordinalzahlen eine lineare Ordnung. Sie ist ferner eine Wohlordnung, da für Mengen $a$ von Ordinalzahlen die Minimalitätsbedingung $\emptyset \neq a \rightarrow \exists \alpha \in a\ \alpha \cap a = \emptyset$ nach unserer Vereinbarung (also dem Fundierungsaxiom) erfüllt ist.

(ii) Angenommen, es gäbe eine Menge $a$ aller Ordinalzahlen. Nach Satz 5.1.3 (iii) ist $a$ transitiv und nach der gerade bewiesenen Aussage (i) ist die $\in$-Beziehung eine Wohlordnung auf $a$, also ist $a$ selbst eine Ordinalzahl, somit $a \in a$ nach Definition von $a$ als Menge aller Ordinalzahlen, aber anderseits gilt $a \notin a$ für alle Ordinalzahlen $a$, Widerspruch! (Man könnte auch so argumentieren: die Menge $a$ aller Ordinalzahlen wäre als Ordinalzahl die größte Ordinalzahl, dann kann aber nicht $a \in a$ sein!) $\square$

Die Antinomie von BURALI-FORTI (1897) war CANTOR übrigens bereits schon 1895 bekannt. Welche Bedeutung hat sie? Sie besagt, daß es keine größte Ordinalzahl gibt, und wir können sie als Aussage verstehen, daß die Gesamtheit aller Ordinalzahlen so "groß" ist, daß sie sich nicht zu einer Menge zusammenfassen läßt. Das wäre an sich harmlos, wenn man nun nicht befürchten müßte, daß vielleicht an einer anderen Stelle der Theorie ein (womöglich bisher noch gar nicht entdeckter) Widerspruch versteckt ist, der sich nicht so einfach hinweginterpretieren läßt. Wir wollen daher erst einmal innehalten, um uns den Grundlagen der Mengenlehre zuzuwenden.

## 5.2 Mengen und Klassen

In seinem Hauptwerk[1] gab CANTOR seine berühmte Beschreibung einer Menge:

> *Unter einer "Menge" verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die "Elemente" von M genannt werden) zu einem Ganzen.*

Man kann diese "Definition" in folgendem Sinne zu präzisieren versuchen:

[1] CANTOR, GEORG *Beiträge zu Begründung der transfiniten Mengenlehre.* Math. Ann. 46, (1895), 481 - 512, Math. Ann. 49, (1897), 207 - 246

*Zu jeder Eigenschaft $E(x)$ existiert die Menge $M$ aller Objekte $x$ mit der Eigenschaft $E(x)$: $M = \{x : E(x)\}$.*

ZERMELOs Erklärung des Begriffes Eigenschaft als definite Eigenschaft wurde von FRAENKEL und SKOLEM dahingehend präzisiert, daß Eigenschaften als Ausdrücke einer bestimmten formalen Sprache festzulegen sind. Für diese Sprache, $\mathcal{L}_{ZF}$, reicht es aus, die prädikatenlogische Sprache mit $\in$ als 2-stelligem Symbol für die Element-Beziehung zu wählen. Die Formeln von $\mathcal{L}_{ZF}$ werden in diesem Teil *mengentheoretische Formeln* (oder einfach: *Formeln*) genannt und wie früher mit $\varphi, \psi, \ldots$ bezeichnet. Die CANTORsche Definition kann nun als Forderung präzisiert werden:

### 5.2.1 Komprehensionsaxiom

$$\exists y \forall x (x \in y \leftrightarrow \varphi(x))$$

*wobei $\varphi(x)$ eine beliebige Formel ist, in welcher $y$ nicht vorkommt.*

Dieses Axiom (eigentlich ein Axiomenschema, da es aus unendlich vielen Axiomen besteht) ist nun aber widerspruchsvoll: Wie wir gerade mittels der Antinomie von BURALI-FORTI gesehen haben, führt es für die Eigenschaft $Ord(x)$ zu Widerspruch, da es keine Menge aller Ordinalzahlen geben kann. B. RUSSELL entdeckte einen Widerspruch, der viel elementarer zu erhalten ist:

**RUSSELLsche Antinomie (1903)**

Wählt man für $\varphi(x)$ die Formel $x \notin x$, so liefert das Komprehensionsaxiom die Existenz einer Menge $r$ mit

$$\begin{array}{rl} \forall x (x \in r \leftrightarrow x \notin x) & \text{insbesondere gilt für } x = r: \\ r \in r \leftrightarrow r \notin r & \qquad \text{Widerspruch!} \end{array}$$

Ähnliche Widersprüche hatte man zwar schon früher erhalten (Antinomie der “Menge aller Mengen” von CANTOR, Antinomie der “Menge aller Ordinalzahlen” von CANTOR und BURALI-FORTI). Diese frühen Antinomien waren jedoch weniger beunruhigend, eher konnten sie positiv bewertet werden als Ausdruck der Reichhaltigkeit des Mengen- und Ordinalzahlbegriffes sowie als Ausweis einer Grenze, über die man nicht hinausgehen konnte. Dagegen war die Antinomie von RUSSELL (die um die gleiche Zeit auch bereits ZERMELO bekannt war) besonders elementar und ließ sich auch umschreiben in einen Widerspruch innerhalb der Prädikatenlogik, wie sie FREGE in seiner *Begriffsschrift* (1879) zu präzisieren versucht hatte. Als 1903 der 2. Band seiner *Grundgesetze der Arithmetik* im Druck war, teilte RUSSELL ihm seine Antinomie brieflich mit, und Frege erklärte daraufhin eine der Grundlagen seiner Theorie als erschüttert.

### 5.2.2 Auswege aus den Antinomien

a) RUSSELLsche Typentheorie

   Im Komprehensionsaxiom darf zur Vermeidung eines circulus vitiosus bei der Definition der Menge $y$ durch die Eigenschaft $\varphi(x)$ nicht auf einen Bereich Bezug genommen werden, dem dieses $y$ selbst angehört. Man nimmt eine Stufeneinteilung vor: $x^0, x^1, \ldots$ und vereinbart, daß $x^n \in x^m$ nur erlaubt ist, wenn $m = n+1$. Das Komprehensionsaxiom hat dann die Form

$$\exists y^{n+1} \forall x^n (x^n \in y^{n+1} \leftrightarrow \varphi(x^n)).$$

   Da die Bildung von $r^n \notin r^n$ nicht mehr erlaubt ist, ist die RUSSELLsche Antinomie nicht mehr formulierbar. Dieses *prädikative* Komprehensionsaxiom schränkt die Mengenbildung jedoch stark ein; die Idee eines stufenweisen Aufbaus der Mengenlehre läßt sich jedoch auch in der ZF-Mengenlehre wieder finden.

b) von NEUMANN, GÖDEL, BERNAYS: Unterscheidung zwischen Mengen und Klassen

   Man läßt eine (mehr oder weniger beschränkte) Komprehension von Mengen zu, das Ergebnis ist dann aber nicht notwendig wieder eine Menge, sondern zunächst eine **Klasse** $\{x \mid \varphi(x)\}$. So kann man bilden:

   *die* RUSSELL*sche Klasse* $R = \{x \mid x \notin x\}$,
   *die Klasse aller Mengen* $V = \{x \mid x = x\}$,
   *die Klasse aller Ordinalzahlen, Kardinalzahlen, usw.*

   Das Auftreten von Antinomien wird nun dadurch vermieden, daß Klassen nicht notwendig wieder Mengen sind; z. B. kann $R$ in die Eigenschaft, welche $R$ definiert, nur dann eingesetzt werden, wenn $R$ eine Menge $x$ ist, und da diese Annahme zum Widerspruch führt, kann $R$ keine Menge sein, sondern nur eine Klasse (**echte Klasse** oder **Unmenge**). Welche Klassen zugleich auch Mengen sind, versucht man durch Axiome festzulegen, wobei man

- einerseits möglichst viele Klassen als Mengen (und damit wieder als Elemente von Mengen und Klassen) erlauben möchte,
- andererseits aber nicht so viele, daß es zu einem Widerspruch kommt.

c) ZERMELOsches Aussonderungsaxiom

Das Komprehensionsaxiom wird eingeschränkt auf die Bildung von Teilmengen einer bereits gegebenen Menge $a$:

**Aussonderungsaxiom:** $\exists y \forall x(x \in y \leftrightarrow x \in a \wedge \varphi(x))$

Versucht man erneut, die RUSSELLsche Antinomie abzuleiten, so erhält man die Existenz einer Menge $r$ mit

$$r \in r \leftarrow r \in a \wedge r \notin r, \text{ also}$$

- es gibt keine "Menge aller Mengen" (für $a$ eingesetzt, erhielte man den RUSSELLschen Widerspruch),
- es gibt eine "leere" Menge (setze für $\varphi(x)$ ein: $x \neq x$ und für $a$ irgendeine Menge).

Das Aussonderungsaxiom ist nachweisbar widerspruchsfrei - es ist bereits wahr in einem Modell, welches nur die leere Menge enthält. Um es darüber hinaus anwenden zu können, müssen wir die Menge $a$, aus welcher mittels der Eigenschaft $\varphi(x)$ eine Teilmenge *ausgesondert* wird, bereits haben - weitere Axiome sind erforderlich! Im folgenden werden wir - wie jetzt allgemein üblich - der Entwicklung der Mengenlehre nach ZERMELO folgen, dabei aber auch den Klassenbegriff benutzen, weil sich damit die Axiome von ZF leicht in der Form

*bestimmte Klassen (und zwar "nicht zu große") sind Mengen*

bequem ausdrücken lassen.

### 5.2.3 Die mengentheoretische Sprache mit Klassentermen

Wie schon angedeutet, benötigen wir für die ZF-Sprache nur die $\in$-Beziehung als einziges nicht-logisches Prädikat. Zur Abwechslung wollen wir hier syntaktisch zwischen freien und gebundenen Variablen unterscheiden: Als *Variable für Mengen* benutzen wir Kleinbuchstaben, und zwar meistens

- $a, b, c, \ldots$ für *freie* Variable und

- $x, y, z, \ldots$ (u. U. mit Indizes) für *gebundene* Variable.

Die damit gebildete Sprache $\mathcal{L}_{\mathsf{ZF}}$ werden wir erweitern um die Hinzunahme von definierbaren Klassen (auch *Klassenterme* genannt)

$$\{x \mid \varphi(x)\},$$

welche in Formeln jedoch nur in der Kombination $a \in \{x \mid \varphi(x)\}$ benötigt werden, so daß **Formeln** wie folgt gebildet werden:

(F1) Sind $a, b$ Mengenvariable, so sind $a = b$ und $a \in b$ Formeln (*Primformeln, atomare Formeln*).

(F2) Ist $\varphi$ eine Formel, so auch $\neg\varphi$.

(F3) Sind $\varphi$ und $\psi$ Formeln, so auch $(\varphi \vee \psi)$.

(F4) Ist $\varphi(a)$ eine Formel, in welcher die Variable $x$ nicht vorkommt, so auch $\exists x \varphi(x)$.

(F5) Ist $\varphi(a)$ eine Formel, in welcher die Variable $x$ nicht vorkommt, so ist auch $a \in \{x \mid \varphi(x)\}$ eine Formel.

(F6) Das sind alle Formeln.

Von Formeln der ZF-Sprache im **engeren** Sinne sprechen wir, wenn bei der Bildung dieser Formeln die Bedingung (F5) nicht benutzt wurde, ansonsten von Formeln der **erweiterten** ZF-Sprache.

Die logischen Symbole $\wedge, \rightarrow, \leftrightarrow$ sowie der $\forall$-Quantor werden wie früher definiert. Außerdem benutzen wir **beschränkte Quantoren** $\forall x \in a$, $\exists x \in a$ als Abkürzungen:

$$\begin{array}{lll} \forall x \in a\, \varphi & \text{steht für} & \forall x(x \in a \rightarrow \varphi), \\ \exists x \in a\, \varphi & \text{steht für} & \exists x(x \in a \wedge \varphi). \end{array}$$

Für viele Untersuchungen spielen Fragen der Komplexität von Eigenschaften eine wichtige Rolle; dabei gelten beschränkte Quantoren als "einfacher" als unbeschränkte Quantoren. Wir werden daher im folgenden - soweit möglich - beschränkte Quantoren verwenden.

### Elimination von Klassentermen

Unser erstes Axiom (für die erweiterte ZF-Sprache) ist das folgende

**CHURCHsche Schema** (CS): $a \in \{x \mid \varphi(x)\} \leftrightarrow \varphi(a)$.

Es erlaubt, Klassenterme wieder zu eliminieren.

Die erweiterte ZF-Sprache ist somit eigentlich nicht ausdrucksstärker, aber es wird sich zeigen, daß sie für mengentheoretische Überlegungen bequemer zu benutzen ist, vor allem weil man für jede Eigenschaft $\varphi(x,\ldots)$ die Klasse $\{x \mid \varphi(x,\ldots)\}$ bilden kann, während man zur Bildung einer entsprechenden Menge erst nachprüfen muß, ob die Axiome dies zulassen bzw. ob man einen Widerspruch erhält, was sehr schwierig (oder sogar unmöglich) sein kann!

### Variable für Klassen

Schließlich noch ein weiterer Schritt: Um nicht immer über die Klassenterme $\{x \mid \varphi(x,\ldots)\}$ durch Angabe von Formeln sprechen zu müssen, werden neue Variable $A, B, \ldots$ für Klassenterme benutzt. Damit nähern wir uns einer Sprache 2. Stufe, in welcher neben den Variablen für Mengen (Kleinbuchstaben) auch Variable für Klassen (Großbuchstaben) auftreten (vgl. die entsprechende Situation in der PEANO-Arithmetik 3.5.4). Streng genommen werden hier keine formalen Klassenvariable eingeführt, sondern **Metavariable** für Klassenterme. Das bedeutet, daß eine Aussage der Form

$$\textit{es gibt ein } A \textit{ mit} \ldots A \ldots B \ldots$$

als Behauptung zu lesen ist, für einen vorgegebenen Klassenterm $B$ einen Klassenterm $A$ konkret (d. h. durch eine Formel) anzugeben, der die gewünschte Eigenschaft besitzt; ebenso bedeutet eine universelle Aussage der Form

$$\textit{für alle } A \ldots A \ldots$$

daß sie für alle (konkret definierbaren) Klassenterme $A$ gilt.

Es gibt nun allerdings Axiomensysteme der Mengenlehre mit freien und möglicherweise auch gebundenen Klassenvariable (also in einer Sprache 2. Stufe mit formalen Variablen wie den Mengenvariablen), in denen Klassen eine stärkere Rolle spielen:

### 5.2.4 Überblick über verschiedene Axiomensysteme

Die Axiomensysteme der Mengenlehre unterscheiden sich weniger in dem jeweils zugelassenen Bereich von Mengen als in dem Status, den sie den Klassen einräumen:

- ohne Klassen:
  das Axiomensystem ZF von ZERMELO-FRAENKEL in der (engeren) ZF-Sprache

- mit eliminierbaren Klassen ("virtuelle Klassen" (QUINE)):
  das Axiomensystem ZF von ZERMELO-FRAENKEL in der erweiterten ZF-Sprache (welches wir hier zugrunde legen werden)

- mit freien Klassenvariablen:
  das Axiomensystem von BERNAYS 1958

- mit freien und gebundenen Klassenvariablen, aber nur prädikativen Klassen:
  das Axiomensystem NBG (von NEUMANN-GÖDEL-BERNAYS)) mit dem Komprehensionsaxiom in der Form

  $$\exists Y \forall x (x \in Y \leftrightarrow \varphi(x, \dots, A, \dots)).$$

  Hier stehen $A, B, \dots, X, Y, \dots$ für Klassen, $x, y, \dots$ weiterhin für Mengen, und $\varphi(x, \dots)$ ist eine Eigenschaft von Mengen (möglicherweise mit Klassen als Parametern), in welcher aber nur über *Mengen* quantifiziert werden darf - prädikative Form).

- Klassen, ernst genommen (imprädikativ):
  das Axiomensystem QM von QUINE-MORSE mit dem Komprehensionsaxiom wie oben, aber jetzt werden auch Formeln zugelassen, die Quantifikationen über Klassen enthalten; in diesem System lassen sich mehr Aussagen über Mengen beweisen als in NBG!

Neben der Entscheidung, welchen existentiellen Wert man den Klassen einräumen soll, bleibt als Hauptproblem die Frage: welche *Mengen* existieren - bzw.:

(i) welche Eigenschaften definieren Mengen - oder

(ii) (wenn wir ein allgemeines Komprehensionsprinzip für Klassen zugrunde legen): welche Klassen führen zu Mengen?

Grundlegend werden folgende Vorstellungen sein:

- Es gibt *Mengen* und *Klassen*; ihre Gleichheit wird durch das Extensionalitätsprinzip bestimmt: Mengen bzw. Klassen sind identisch, wenn sie vom selben Umfang sind, d. h. dieselben Elemente enthalten.

- *Elemente* von Klassen und Mengen sind wieder *Mengen*; echte Klassen (wie die RUSSELLsche Klasse) sind dagegen niemals Elemente - weder von Mengen noch von anderen Klassen.

- Komprehensionsprinzip: Jeder Eigenschaft $E(x)$ von Mengen $x$ entspricht die *Klasse* $\{x \mid E(x)\}$ aller *Mengen* $x$ mit der Eigenschaft $E(x)$,

- insbesondere ist jede Menge $a$ eine Klasse (nämlich: $a = \{x \mid x \in a\}$),

- umgekehrt ist aber nicht jede Klasse eine Menge, sondern nur diejenigen Klassen, die nicht "zu groß" sind, sind Mengen (*limitation of size*).

*Klassen* sind somit nach dem Komprehensionsprinzip (nahezu) unbeschränkt bildbar, *Mengen* existieren (als spezielle Klassen) nur nach Maßgabe einzelner Axiome.

**Literatur zu diesem Kapitel**

BERNAYS, P. *Axiomatic Set Theory.* Amsterdam 1958

FELGNER, U. (Herausgeber) *Mengenlehre.* Darmstadt 1979

FRAENKEL, A.- BAR HILLEL, Y. - LÉVY, A. *Foundations of Set Theory.* Amsterdam 1973

FRAENKEL, A.A. *Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre.* Math Ann 86 (1922), 230 - 237, auch in: FELGNER [1979]

SKOLEM, T. *Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre.* Kongreß Helsingfors 1922 (auch in FELGNER [1979])

## 5.3 Extensionalität und Aussonderung

Wie bereits erwähnt, sind Mengen bestimmt allein durch ihre *Elemente* (wobei die Elemente ihrerseits natürlich wieder Mengen sind). Diese Vorstellung wird ausgedrückt durch das

**Extensionalitätsaxiom** (Ext) $\forall x(x \in a \leftrightarrow x \in b) \to a = b.$

Aufgrund der logischen Axiome über die Gleichheit gilt auch die Umkehrung, also:

$$a = b \leftrightarrow \forall x(x \in a \leftrightarrow x \in b).$$

Somit könnte man im Prinzip auf die Gleichheit als logisches Grundsymbol verzichten und sie mittels der Elementbeziehung definieren. - Denkbar wäre auch eine Definition durch

**intensionale (Leibniz-)Gleichheit:** $\forall x(a \in x \leftrightarrow b \in x) \leftrightarrow a = b.$

Diese Äquivalenz wird später einfach beweisbar sein (mit Existenz der Einer-Menge).

Mengen ohne Elemente bezeichnet man auch als **Urelemente**; aus dem Extensionalitätsaxiom folgt, daß es höchstens eine solche Menge gibt, die leere Menge. Es besteht durchaus die Möglichkeit, weitere derartige Mengen als *Urelemente* zuzulassen, für den mengentheoretischen Aufbau der Mathematik werden sie jedoch nicht benötigt.

Die Gleichheit von Klassen bzw. Klassen und Mengen wird als Umfangsgleichheit <u>definiert</u>:

$$\begin{aligned} A = B :\ &\leftrightarrow\ \forall x(x \in A \leftrightarrow x \in B), \\ A = b :\ &\leftrightarrow\ \forall x(x \in A \leftrightarrow x \in b), \\ a = B :\ &\leftrightarrow\ \forall x(x \in a \leftrightarrow x \in B), \end{aligned}$$

und in ähnlicher Weise (da Elemente stets Mengen sein sollen):

$$\begin{aligned} A \in B :\ &\leftrightarrow\ \exists y(y = A \wedge y \in B), \\ A \in b :\ &\leftrightarrow\ \exists y(y = A \wedge y \in b). \end{aligned}$$

Somit sind nun die Grundprädikate $=$ und $\in$ für sowohl Mengen wie auch Klassen definiert.

Wenn eine Klasse dieselben Elemente wie eine Menge enthält, so sind nach unserer Definition beide gleich, und genau in diesem Fall werden wir eine Klasse eine Menge nennen:

| | |
|---|---|
| $Mg(A) :\leftrightarrow \exists x(x = A)$ | A ist eine **Menge**, somit gilt also |
| $Mg(A) \leftrightarrow A \in V,$ | wobei |
| $V := \{x \mid x = x\}$ | **Klasse aller Mengen, Allklasse, Universum**. |

(Klassen, die keine Mengen sind, nennt man auch **echte** Klassen.)

Von der Element-Beziehung zu unterscheiden ist die *Teilmengen-Beziehung* (**Inklusion**) (auch wenn man in beiden Fällen vom "Enthalten" spricht):

$$A \subseteq B : \quad \leftrightarrow \quad \forall x(x \in A \rightarrow x \in B)$$
$$A \subset B : \quad \leftrightarrow \quad A \subseteq B \wedge A \neq B$$

Somit kann man die Gleichheit zweier Mengen nachweisen, indem man zeigt, daß sie wechselweise ineinander enthalten sind:

$$A = B \leftrightarrow A \subseteq B \wedge B \subseteq A.$$

**Die Booleschen Operationen**

können wir allein mit Hilfe der Klassenschreibweise einführen:

| | |
|---|---|
| $A \cap B := \{x \mid x \in A \wedge x \in B\}$ | **Durchschnitt** |
| $A \cup B := \{x \mid x \in A \vee x \in B\}$ | **Vereinigung** |
| $A - B := \{x \mid x \in A \wedge x \notin B\}$ | **relatives Komplement** |
| $-B := \{x \mid x \notin B\}$ | **Komplement** |

Damit erhalten wir nun die Gesetze einer BOOLEschen Algebra, wie wir sie bereits von der Aussagenlogik her kennen (s. 1.2.6). Sie gelten statt für Klassen natürlich auch für Mengen, allerdings führen nicht alle Operationen von Mengen zu Mengen (aber von Klassen zu Klassen), denn in den üblichen mengentheoretischen Axiomensystemen ist $V$ (und das Komplement einer Menge) keine Menge. Bereits erwähnt haben wir das

**ZERMELOsche Aussonderungsschema** (AusS)

$$\exists y \forall x(x \in y \leftrightarrow x \in a \wedge \varphi(x, \ldots))$$

Unter Benutzung von Klassenvariablen und einfachen Definitionen erhalten wir den

**Satz**

*Das Aussonderungsschema ist äquivalent zu den folgenden Aussagen:*

(i) $\exists y \forall x(x \in y \leftrightarrow x \in a \wedge x \in A)$ **Aussonderungsaxiom**

(ii) $\exists y(y = a \cap A)$

(iii) $a \cap A \in V$, *d. h.* $Mg(a \cap A)$

(iv) $A \subseteq a \to Mg(A)$ **Abschätzungssatz** $\square$

Somit erhalten wir die Existenz folgender Mengen:

$$\emptyset,\ a \cap b,\ a - b,$$

wobei $b$ auch eine Klasse $B$ sein darf. Dagegen kann die Allklasse $V$ keine Menge sein.

## 5.4 Relationen und Funktionen

Um Relationen und Funktionen als spezielle Klassen einzuführen, benutzen wir geordnete Paare: Zunächst können wir folgende Klassen bilden:

$\{a,b\} := \{x \mid x = a \vee x = b\}$ **ungeordnetes Paar**
$\{a\} := \{x \mid x = a\} = \{a,a\}$ **Einerklasse von a**.

Dieses sind offensichtlich sehr kleine Klassen, wir können also das

**Paarmengenaxiom** (Paar) $\exists y \forall x(x \in y \leftrightarrow x = a \vee x = b)$, d. h. $Mg(\{a,b\})$

annehmen und damit die folgenden Mengen erhalten:

$$\{a,b\}, \{a\}, \{a,\{a,b\}\}, \{\{a,c\},\{e,\{a,e\}\}\}, \{\{a,\{a,b\}\}, \{\{a,c\},\{e,\{a,e\}\}\}, \ldots \text{ usw.}$$

**Geordnetes Paar (Wiener 1914, Kuratowski 1921)**

$(a,b) := \{\{a\},\{a,b\}\}$ **geordnetes Paar**

Mittels geordneter Paare lassen sich auch **Tripel** $(a,b,c) := (a,(b,c))$ und allgemeiner $n$-Tupel definieren. Die genaue Definition des geordneten Paares und des $n$-Tupels ist hier nicht wichtig, wir benötigen nur ihre charakteristischen Eigenschaften:

$$(a,b) = (c,d) \ \leftrightarrow \ a = c \ \wedge \ b = d, \quad \text{und allgemein:}$$
$$(a_1,\dots,a_n) = (b_1,\dots,b_n) \leftrightarrow a_i = b_i \ \textit{für alle } i = 1,\dots,n.$$

Damit können wir damit 2-stellige Relationen (und ähnlich mittels $n$-Tupeln $n$-stellige Relationen) auf Klassen zurückführen:

$$(a,b) \in A \leftrightarrow \varphi(a,b),$$

entsprechend unserer früheren Einführung von Klassen als Extensionen 1-stelliger Prädikate:

$$a \in A \leftrightarrow \varphi(a).$$

Die **Klasse der Paare** $(x,y)$ mit der Eigenschaft $\varphi(x,y)$ ist

$$\{x,y \mid \varphi(x,y)\} := \{z \mid \exists x \exists y (z = (x,y) \wedge \varphi(x,y))\}$$
$$\text{oder auch geschrieben } \{(x,y) \mid \varphi(x,y)\}.$$

Sonderfälle sind

$A \times B \ := \{(x,y) \mid x \in A \wedge \ y \in B\},$ **cartesisches Produkt**,
$A^n \ := \{(x_1,\dots,x_n) \mid x_1,\dots,x_n \in A\}.$ n-faches **Produkt** von $A$.

$n$-stellige Relationen identifizieren wir mit den $n$-Tupeln, die diese Relationen erfüllen:

$$\boxed{R \textit{ ist } n\text{-stellige } \textbf{Relation} \text{ auf } A :\leftrightarrow R \subseteq A^n}$$

Im Falle einer 2-stelligen Relation schreiben wir dann auch wie üblich:

$$aRb :\leftrightarrow (a,b) \in R.$$

**Funktionen** sind 2-stellige Relationen von Paaren $(x,y)$, in denen $y$ durch $x$ eindeutig bestimmt ist:

$$Fkt(F) :\leftrightarrow F \subseteq V \times V \wedge \forall x,y,z\, ((x,y) \in F \wedge (x,z) \in F \rightarrow y = z).$$

Somit werden Funktionen wir mit ihren *Graphen* identifiziert:

$$\boxed{F : A \longrightarrow B :\leftrightarrow\ F \subseteq A \times B\ \wedge\ \forall x \in A\ \exists! y \in B\ (x,y) \in F}$$

Hierfür sagen wir auch: $F$ *ist* **Funktion von** $A$ **nach** $B$. Für $x \in A$ bezeichnet

$$F(x) \text{ das eindeutig bestimmte } y \text{ mit } (x,y) \in F$$

den **Funktionswert** von $F$ an der Stelle $x$, und für $D \subseteq A$ ist

$$\begin{aligned} \{F(x) \mid x \in D\} &:= F[D] := \{y \mid \exists x(x \in D \wedge y = F(x))\} \text{ das } F\text{-}\textbf{Bild von } D \\ F \restriction D &:= \{(x, F(x)) \mid x \in D\} \text{ die Funktion } F, \textbf{ eingeschränkt auf } D. \end{aligned}$$

Im Falle von Strukturen werden wir Funktionen $F$ mit $F : A^n \to A$ betrachten; wir sagen dann auch, daß $F$ eine $n$-stellige Funktion **auf der Menge** $A$ ist.

Häufig definiert man eine Funktion durch Angabe ihres Wertes $F(x)$ für $x \in A$:

$$\begin{aligned} F : A &\longrightarrow B, \\ x &\mapsto F(x). \end{aligned}$$

Unter den Funktionsbegriff fällt auch der Begriff der **Familie**: Für eine Indexmenge $I$ ist

$$(x_i)_{i \in I} \text{ die Funktion } f \text{ mit Definitionsbereich } I \text{ und } f(i) = x_i \text{ für alle } i \in I.$$

Im Falle $I = \mathbb{N}$ spricht man auch von einer **Folge**, für $I = \{0, \dots, n-1\}$ entspricht die endliche Folge einem $n$-Tupel.

### Injektive, surjektive und bijektive Funktionen

Ist $F : A \longrightarrow B$, also $F$ eine Abbildung von $A$ **nach** $B$, so nennen wir

- $A$ **Definitionsbereich** von $F$,
- $B$ **Bildbereich** von $F$, und
- $W(F) := \{F(x) \mid x \in A\}$ **Wertebereich** von $F$.

(*Bild-* und *Wertebereich* werden oft gerade in umgekehrter Weise bezeichnet!)

Der Bildbereich $B$ ist durch $F$ nicht eindeutig bestimmt; er braucht nämlich den Wertebereich $W(F)$ nur zu enthalten. Ist jedoch Bildbereich = Wertebereich, so heißt $F$ **surjektiv** (oder: **auf** $B$):

$$F : A \twoheadrightarrow B :\leftrightarrow F : A \longrightarrow B\ \wedge\ B = \{F(x) \mid x \in A\}.$$

In diesem Fall gibt es also zu jedem $b \in B$ *mindestens* ein $a \in A$ mit $F(a) = b$. Die Angabe des Bildbereiches $B$ ist in diesem Fall wichtig, da jede Funktion eine Abbildung *auf* ihren Wertebereich ist!

Eine Funktion $F : A \longrightarrow B$ heißt **injektiv** oder **eineindeutig** gdw es zu jedem $b \in W(F)$ *genau* ein $a \in A$ mit $F(a) = b$ gibt, welches mit $a = F^{-1}(b)$ als **Urbild** von $b$ bezeichnet wird:

$$F : A \rightarrowtail B :\leftrightarrow F : A \longrightarrow B \wedge \forall x, y \in A \, (F(x) = F(y) \rightarrow x = y).$$

Für eine injektive $F$ Funktion gibt es also eine inverse Funktion $F^{-1}$, und zwar ist

$$F^{-1} : W(F) \longrightarrow A \quad \text{für injektives } F : A \longrightarrow B.$$

Eine Funktion $F : A \longrightarrow B$ heißt **bijektiv** gdw $F$ injektiv und surjektiv ist:

$$F : A \longleftrightarrow B :\leftrightarrow F : A \rightarrowtail B \wedge F : A \twoheadrightarrow B.$$

Für eine bijektive Funktion $F : A \longleftrightarrow B$ ist dann die **inverse Funktion** (oder **Umkehrfunktion**) eine (ebenfalls bijektive) Abbildung

$$\begin{aligned} F^{-1} : B \longleftrightarrow A \quad & \text{mit} \\ F^{-1}(F(a)) = a,\ F(F^{-1}(b)) = b \quad & \text{für alle } a \in A, b \in B. \end{aligned}$$

## 5.5 Vereinigung und Produkt

### 5.5.1 Vereinigung

Eine einfache mengentheoretische Operation führt von zwei Mengen zu ihrer Vereinigung; wir werden gleich eine allgemeinere Mengenbildung einführen, nämlich die Vereinigung der Elemente einer Klasse (und dual der entsprechende Durchschnitt):

$$\begin{array}{ll} \bigcup A := \bigcup_{x \in A} x := \{z \mid \exists x \in A \; z \in x\} & \textbf{Vereinigung (Summe)} \\ \bigcap A := \bigcap_{x \in A} x := \{z \mid \forall x \in A \; z \in x\} & \textbf{Durchschnitt} \end{array}$$

Die Vereinigung der Elemente einer Menge (welches stets Mengen sind) sollte wieder eine Menge ergeben:

**Summenaxiom (Zermelo)** (Sum) $\quad Mg(\bigcup a)$,
ausgeschrieben also: $\quad \exists y \forall z (z \in y \leftrightarrow \exists x \in a \; z \in x)$.

**Lemma**

(i) $\bigcup\{a\} = a, \quad \bigcup\{a,b\} = a \cup b$

(ii) $\bigcap\{a\} = a, \quad \bigcap\{a,b\} = a \cap b$

(iii) $\bigcup a$ *ist das Supremum der Elemente von a bzgl.* $\subseteq$:
$\forall x \in a\, x \subseteq \bigcup a, \forall x \in a\, x \subseteq c \to \bigcup a \subseteq c.$

(iv) $\bigcap a$ *ist das Infimum der Elemente von a bzgl.* $\subseteq$:
$\forall x \in a \bigcup a \subseteq x, \forall x \in a\, c \subseteq x \to c \subseteq \bigcup a.$

(v) $\bigcup \emptyset = \emptyset, \quad \bigcup V = V,$
$\bigcap \emptyset = V, \quad \bigcap V = \emptyset\,.$

**Bemerkungen**

1. Aus dem Summen- und Paarmengenaxiom folgt also, daß für je zwei Mengen $a$ und $b$ die Vereinigung $a \cup b$ wieder eine Menge ist, damit ist dann aber das absolute Komplement einer Menge $-a$ stets eine echte Klasse!

2. Dagegen ist $a \cap b$ eine Menge nach dem Aussonderungsaxiom, allgemeiner gilt nach (AusS):

$$A \neq \emptyset \to Mg(\bigcap A), \quad \text{wegen } \bigcap \emptyset = V \text{ also}$$
$$A \neq \emptyset \leftrightarrow Mg(\bigcap A).$$

Statt der obigen Summen und Durchschnitte kommen in der Mathematik häufiger Vereinigungen und Durchschnitte von Familien von Mengen vor. Zur Angleichung an den dort üblichen Gebrauch bezeichnen wir den Indexbereich mit I (der zunächst eine Klasse sein kann) und schreiben auch $F_i$ für $F(i)$:

$\bigcup_{i \in I} F_i := \bigcup\{F_i \mid i \in I\} = \{z \mid \exists i \in I\, z \in F_i\}$ **Vereinigung**
$\bigcap_{i \in I} F_i := \bigcap\{F_i \mid i \in I\} = \{z \mid \forall i \in I\, z \in F_i\}$ **Durchschnitt**

In Verallgemeinerung des ZERMELOschen Axioms besagt das Summenaxiom von BERNAYS, daß die Vereinigung von Mengen-vielen Mengen wieder eine Menge ist:

**BERNAYSsches Summenaxiom** (BSum) $\quad Mg(\bigcup_{x \in a} F(x))$

Der Indexbereich $I$ ist hier eine Menge $a$, jedem $i \in a$ ist eine *Menge* $F(i)$ zugeordnet, die Zuordnung kann allerdings durch eine Klasse $F$ gegeben sein! Wenn man dies beachtet, kann man die BERNAYSsche Summe auch in der vertrauten Form $\bigcup_{i \in I} a_i$ und das BERNAYSsche Summenaxiom in der Form

$$Mg(\bigcup_{i \in I} a_i) \text{ für jede Indexmenge } I$$

schreiben. Als Axiom wählen statt dessen das

**Ersetzungsaxiom** (Ers) $\qquad Mg(\{F(x) \mid x \in a\})$

von A. FRAENKEL. Es besagt, daß das Bild einer *Menge* $a$ unter einer Funktion $F$ (welche eine *Klasse* sein kann) wieder eine Menge ist, bzw. daß man wieder eine Menge erhält, wenn man die Elemente $x$ einer Menge $a$ durch ihre Funktionswerte $F(x)$ *ersetzt*. Wir können es auch in der folgenden Form schreiben:

$$Fkt(F) \to Mg(F[a]) \qquad \text{bzw.}$$
$$Fkt(F) \wedge Mg(D(F)) \to Mg(W(F)).$$

**Satz**

*Auf der Basis der bisherigen Axiome* (Ext, Null, Paar) *gilt:*

BERNAYSsches Summenaxiom = Summenaxiom + Ersetzungsaxiom.

*Beweis:* Aus (BSum) folgt (Sum), indem man für $F$ die Identität (also $F(x) = x$) setzt, und das (Ers) ergibt sich wegen $\{F(x) \mid x \in a\} = \bigcup\{\{F(x)\} \mid x \in a\}$. Umgekehrt gilt:

$$\bigcup_{x \in a} F(x) = \bigcup\{F(x) \mid x \in a\} = \bigcup F[a].$$

□

**Satz**

(i) $Fkt(F) \wedge D(F) \in V \to W(F),\ F \in V$,
d. h. *ist der Definitionsbereich einer Funktion eine Menge, so ist nicht nur ihr Bild, sondern die Funktion selber eine Menge.*

(ii) $F : A \twoheadrightarrow B \wedge Mg(A) \to Mg(B)$

(iii) $F : A \rightarrowtail B \wedge Mg(B) \to Mg(A)$

*Beweis* von (i): Zu gegebener Funktion $F$ definiere eine Funktion $G$ durch

$$D(G) = D(F) \text{ und } G(x) = (x, F(x)) \text{ für } x \in D(F).$$

Dann ist $W(G) = \{G(x) \mid x \in D(F)\} = \{(x, F(x)) \mid x \in D(F)\} = F$, also auch $F$ Menge nach dem Ersetzungsaxiom (angewandt auf $G$). Die Beweise von (ii) und (iii) sind leichte Übungsaufgaben. □

Die Aussagen (ii) und (iii) werden verständlich, wenn man "Mg(A)" inhaltlich deutet als "A ist nicht zu groß" und beachtet, daß das Bild einer Funktion eher "kleiner" als der Definitionsbereich ist (da die Funktion möglicherweise verschiedene Mengen auf dasselbe Objekt abbildet. (Genaueres hierzu bei der Einführung des Begriffes der Kardinalzahl!)

**Satz über das cartesische Produkt**

*Das Produkt zweier Mengen ist wieder eine Menge:* $Mg(a \times b)$

*Beweis:* Offensichtlich kann man die Menge $b$ auf $\{c\} \times b$ abbilden, dieses ist also eine Menge nach dem Ersetzungsaxiom. Somit kann man eine Funktion $F : a \longrightarrow V$ definieren durch $F(x) = \{x\} \times b$.
Hiermit gilt nun $a \times b = \bigcup_{x \in a}(\{x\} \times b) = \bigcup_{x \in a} F(x)$. □

### 5.5.2 Potenzmenge und allgemeines Produkt

| | |
|---|---|
| $\mathcal{P}(A) := \{x \mid x \subseteq A\}$ | **Potenzklasse** von $A$ |
| ${}^{b}A := \{f \mid f : b \longrightarrow A\}$ | **Belegungsklasse** von $b$ nach $A$ |

**Potenzmengenaxiom** (Pot) $\quad Mg(\mathcal{P}(a))$

**Bemerkungen**

- Da Elemente von Klassen stets Mengen sind, ist $\mathcal{P}(A)$ nur als Klasse aller Teil*mengen* von $A$ definierbar, ebenso ist ${}^{b}A$ nur für Mengen $b$ definierbar.

- Es ist stets $\emptyset, a \in \mathcal{P}(a)$.

- Aufgrund der spezifischen Definition des geordneten Paares gilt:

$$a \times b \subseteq \mathcal{P}(\mathcal{P}(a \cup b)).$$

Also kann man auch mit Hilfe des Potenzmengenaxioms (und des Aussonderungsaxioms) zeigen, daß das Produkt von zwei Mengen wieder eine Menge ist.

- Indem man jeder Menge $x \subseteq a$ ihre charakteristische Funktion $c_x$ zuordnet, erhält man eine Bijektion der Potenzklasse $\mathcal{P}(A)$ auf die Belegungsklasse ${}^b2$ mit $2 := \{0,1\} = \{\emptyset, \{\emptyset\}\}$. Falls $a$ eine endliche Menge mit $n$ Elementen ist, so hat also die Potenzmenge von $a$ $2^n$ Elemente (was auch für unendliche Mengen gilt und die Bezeichnung "Potenzmenge" erklärt).

Wie im Falle der Vereinigung gibt es für eine Menge ein einfaches und für eine Familie von Mengen ein allgemeines Produkt; in diesem Fall läßt es sich aber nur für Index*mengen* definieren:

**Allgemeines Produkt**

$$\prod_{x \in a} F(x) := \{f \mid Fkt(f) \wedge D(f) = a \wedge \forall x \in a\, f(x) \in F(x)\}$$

Weniger gebräuchlich ist der Sonderfall

$$\prod a = \prod_{x \in a} x := \{f \mid Fkt(f) \wedge D(f) = a \wedge \forall x \in a\, f(x) \in x\}.$$

**Bemerkung**

Es gibt einfache Bijektionen (für $a \neq b$)

$$\prod_{x \in \{a\}} F(x) \longleftrightarrow F(a), \qquad \prod_{x \in \{a,b\}} F(x) \longleftrightarrow F(a) \times F(b)$$

usw., so daß man das allgemeine Produkt als Verallgemeinerung des endlichen Produktes auffassen kann. Benutzen wir wieder $I$ als Indexmenge und Familien statt Funktionen, so können wir das allgemeine Produkt auch in der folgenden Form schreiben:

$$\text{für Mengen } I: \quad \prod_{i \in I} F_i \qquad \text{bzw.} \qquad \prod_{i \in I} a_i = \{(x_i)_{i \in I} \mid \forall i \in I\, x_i \in a_i\}$$

wobei wir statt $F$ jetzt auch eine Mengenvariable $a$ einsetzen können, da nach dem Ersetzungsaxiom eine Funktion, die auf einer Menge definiert ist, selbst eine Menge ist.

**Satz**

(i) $\prod_{x\,\in a} F(x) \subseteq \mathcal{P}(a \times \bigcup_{x\in a} F(x))$

(ii) ${}^a b = \prod_{x\,\in a} F(x)$ *mit* $F(x) = b$ *konstant für alle* $x \in a$,

(iii) ${}^a b$ *und* $\prod_{x\,\in a} F(x)$ *sind Mengen.* $\square$

Somit können wir die folgenden Mengen bilden:

$$\emptyset,\ \{a\},\ \{a,b\},\ a\cup b,\ a\cap B,\ a-b,$$
$$\bigcup a,\ \bigcup_{x\in a} F(x),\ F[a],\ \{f \mid f : a \longrightarrow b\},\ a\times b,\ \prod_{x\in a} F(x),\ \mathcal{P}(a).$$

Damit erhalten wir die in der Mathematik gebräuchlichen Operationen, die Mengen ergeben, soweit sie auf Mengen angewandt werden - das gilt zwar nicht für das Komplement, aber dieses wird ohnehin in Anwendungen nur relativ zu einer vorgegebenen Grundmenge gebraucht. Es fehlt aber noch die Existenz einer unendlichen Menge (und damit der Zahlbereiche $\mathbb{N},\mathbb{Z},\mathbb{R},\dots$), und hierzu benötigen wir das

**Unendlichkeitsaxiom** (Un) $\quad \exists x(\emptyset \in x \land \forall y(y \in x \to y \cup \{y\} \in x))$

Die kleinste derartige Menge, d. h. der Durchschnitt

$$\mathbb{N} := \bigcap\{x \mid \emptyset \in x \land \forall y(y \in x \to y \cup \{y\} \in x)\}$$

kann als Menge der natürlichen Zahlen gewählt werden. Im Sinne der Ordinalzahltheorie ist es eine Ordinalzahl, und zwar die kleinste unendliche Ordinalzahl, bezeichnet mit $\omega$.

Die bisherigen Axiome legen die Gleichheit von Mengen fest Ext) oder fordern die Existenz von Mengen, die eindeutig beschrieben werden (Paar, Summe, Potenz, Bildmenge, natürliche Zahlen). Von anderer Art ist das

**Fundierungsaxiom** (Fund) $\quad a \neq \emptyset \to \exists x \in a\, x \cap a = \emptyset.$

Dieses schränkt den Mengenbereich ein, indem es "unbequeme" Mengen ausschließt, die man beim gewöhnlichen Aufbau der Mengenlehre nicht benötigt. Wir haben es bei der Entwicklung der Ordinalzahltheorie bereits benutzt und werden eine weitere Anwendung später darstellen.

## 5.6 Überblick über die ZF-Axiome

| | | |
|---|---|---|
| **Extensionalitätsaxiom** | (Ext) | $\forall x(x \in a \leftrightarrow x \in b) \to a = b$ |
| **Nullmengenaxiom** | (Null) | $\exists y \forall x(x \notin y)$ |
| **Paarmengenaxiom** | (Paar) | $\exists y \forall x(x \in y \leftrightarrow x = a \vee x = b)$ |
| **Summenaxiom** | (Sum) | $\exists y \forall z(z \in y \leftrightarrow \exists x \in a\, z \in x)$ |
| **Ersetzungsschema** | (ErsS) | $\forall x \forall y \forall z(\varphi(x,y) \wedge \varphi(x,z) \to y = z)$ |
| | | $\to \exists u \forall y(y \in u \leftrightarrow \exists x \in a \varphi(x,y))$ |
| **Potenzmengenaxiom** | (Pot) | $\exists y \forall z(z \in y \leftrightarrow z \subseteq a)$ |
| **Unendlichkeitsaxiom** | (Un) | $\exists x(\emptyset \in x \wedge \forall y(y \in x \to y \cup \{y\} \in x))$ |
| **Fundierungsaxiom** | (Fund) | $a \neq \emptyset \to \exists x \in a\, \forall y \in x\, y \notin a.$ |

Unter Benutzung von *Klassentermen*, also Ausdrücken der Form $\{x \mid \varphi(x)\}$, bezeichnet mit Großbuchstaben $A, B, \ldots$ und der Definition

$$Mg(A) :\leftrightarrow \exists x(x = A) \quad A \text{ ist Menge}$$

lassen sich die obigen Axiome auch wie folgt kürzer und prägnanter ausdrücken (mit den üblichen Definitionen von Funktionen als Klassen geordneter Paare):

| | | |
|---|---|---|
| **Extensionalitätsaxiom** | (Ext) | $\forall x(x \in a \leftrightarrow x \in b) \to a = b$ |
| **Nullmengenaxiom** | (Null) | $Mg(\emptyset)$ |
| **Paarmengenaxiom** | (Paar) | $Mg(\{a,b\})$ |
| **Summenaxiom** | (Sum) | $Mg(\bigcup a)$ |
| **Ersetzungsaxiom** | (Ers) | $Fkt(F) \to Mg(F[a])$ |
| **Potenzmengenaxiom** | (Pot) | $Mg(\mathcal{P}(a))$ |
| **Unendlichkeitsaxiom** | (Un) | $Mg(\mathbb{N})$ |
| **Fundierungsaxiom** | (Fund)) | $a \neq \emptyset \to \exists x \in a\, x \cap a = \emptyset.$ |

Das Aussonderungsaxiom ist übrigens als Folge des Ersetzungsaxioms überflüssig (wobei man - je nach Art der Formalisierung - noch das Nullmengenaxiom gebraucht):

**Metatheorem**

(i) *Aus den Axiomen* (Ext), (Null), (Paar), (Ers) *folgt* (Aus),

(ii) *aus* (Ers) *folgt* (AusS).

*Beweis* von (i): zu zeigen ist: $a \cap A \in V$.

1. Fall: $a \cap A = \emptyset$ : Dann gilt die Beh. wegen des Nullmengenaxioms.

2. Fall: $a \cap A \neq \emptyset$: Wähle ein $b \in a \cap A$ und definiere eine Funktion

$F : a \twoheadrightarrow a \cap A$ durch

$$F(x) = \begin{cases} x, & \text{falls } x \in a \cap A, \\ b, & \text{sonst} \end{cases}$$

und wende das Ersetzungsaxiom an.

Dasselbe Argument liefert einen Beweis von (ii), benötigt aber weder das Nullmengenaxiom noch das Paarmengenaxiom:

Zu zeigen ist: $\exists y \forall x (x \in y \leftrightarrow x \in a\ \varphi(x))$.

Definiere dazu eine Formel

$$\psi(x,y) :\leftrightarrow (\varphi(x) \wedge x = y).$$

Es gilt dann:

$$\psi(x,y) \wedge \psi(x,z) \rightarrow y = z,$$

also existiert nach (ErsS) ein $b$ mit

$$\begin{array}{ll} \forall y (y \in b \leftrightarrow \exists x \in a\ \psi(x,y)), & \text{d. h.} \\ \forall y (y \in b \leftrightarrow \exists x \in a\ (\varphi(x) \wedge x = y)), & \text{also} \\ \forall y (y \in b \leftrightarrow y \in a \wedge \varphi(y)) & \end{array}$$

$\square$

# Kapitel 6

# Mengen von Mengen von ...

Wenn man eine Aussage über alle natürlichen Zahlen beweisen will, so kann man sie nicht für jede Zahl einzeln nachprüfen, da dies unendlich-viele Schritte erfordern würde. Stattdessen kann man auf das Induktionsprinzip für die natürlichen Zahlen zurückgreifen, allgemeiner im Falle unendlicher Mengen auf eine Wohlordnung dieser Menge. Damit werden die Elemente der Menge in einer Weise geordnet, daß man sie von den kleineren zu den größeren "durchlaufen" kann, was dann genauer im (transfiniten) Induktionsprinzip ausgedrückt wird. Wir wiederholen zunächst den Begriff der Wohlordnung und erweitern ihn zugleich auf den Fall von Klassen:

## 6.1 Induktion und Rekursion

### 6.1.1 Ordnungen auf Klassen

1. Eine (irreflexive) **teilweise Ordnung** auf einer Klasse $A$ ist eine 2-stellige Relation $< \subseteq A \times A$, für die gilt:

   (a) $\forall x \in A \quad x \not< x$ **irreflexiv**,
   (b) $\forall x, y, z \in A \; (x < y \wedge y < z \rightarrow x < z)$ **transitiv**.

2. Eine (irreflexive) **lineare Ordnung** auf $A$ ist eine teilweise Ordnung $<$ auf $A$ mit
   (c) $\forall x, y \in A \; (x < y \vee x = y \vee y < x)$ **vergleichbar**.

3. Eine **Wohlordnung** auf einer Klasse $A$ ist eine lineare Ordnung $<$ auf $A$, welche zusätzlich die folgenden Bedingungen erfüllt:

   (F1) $\forall z(\emptyset \neq z \subseteq A \rightarrow \exists x \in z\, \forall y \in z\, y \not< x)$ **Minimalitätsbedingung**
   (F2) $\forall y \in A \quad Mg(\{x \mid x < y\})$ **Mengenbedingung**

4. $S(a,<) := [< a] := \{x \mid x < a\}$ $<$-**Segment** von a, Menge der $<$-**Vorgänger** von a

**Bemerkungen**

- Die Mengenbedingung (F2) ist nur wichtig für echte Klassen; ist $A$ (und damit auch die Relation $<$ auf $A$) eine *Menge*, so ist (F2) stets erfüllt, da $\{x \mid x < a\} \subseteq A$.

- Eine Relation, die nur die Bedingungen (F1) und (F2) erfüllt (also ohne notwendig eine lineare Ordnung zu sein), heißt *fundiert*.

- Im Falle der $\in$-Beziehung ist $\{x \mid x < a\} = \{x \mid x \in a\} = a$ und somit (F2) erfüllt. Ist $\in_A$ die $\in$-Beziehung auf einer Klasse A, so gilt: $\in_A$ ist fundiert, denn $\in_A$ erfüllt auch die Bedingungen (F1) aufgrund des Fundierungsaxioms. Damit ist $\in_A$ zwar auch irreflexiv, aber i. a. keine Wohlordnung (da weder (b) noch (c) gelten). Trotzdem kann man die wichtigsten Ergebnisse über Wohlordnungen auf fundierte Relationen (wie $\in_A$) übertragen.

- Ist $<$ Wohlordnung auf $A$, so läßt sich die Minimalitätsbedingung (F1) auf Klassen verallgemeinern - als formale Bedingung mit einem Allquantor über alle Klassen $B$ können wir sie nicht aufschreiben, sondern nur als Aussagenschema für alle Klassenterme $B$:

### 6.1.2 Minimumsprinzip

*$<$ sei Wohlordnung auf A. Dann hat jede nicht-leere Teilklasse von A ein $<$-minimales (bzw. $<$-kleinstes) Element:*

(i) $\emptyset \neq B \subseteq A \rightarrow \exists x \in B\, \forall y \in B\, y \not< x$ , *d. h. (da $<$ lineare Ordnung ist):*

(ii) $\emptyset \neq B \subseteq A \rightarrow \exists x \in B\, \forall y \in B\, x \leq y$.

*Beweis:* Sei $b \in B$. Falls $b$ nicht bereits $<$-minimal ist, so

$$\exists y \in B\, y < b, \text{ d. h. } [< b] \cap B \neq \emptyset.$$

Nach der Mengenbedingung (F2) ist $[< b] \cap B$ eine Menge und hat damit nach (F1) ein $<$-minimales $c$:

(+) $\quad c \in [< b] \cap B \wedge \forall y \in [< b] \cap B\, y \not< c.$

Dann gilt aber auch $c \in B \wedge \forall y \in B\, y \not< c$,
denn gäbe es ein $y \in B$ mit $y < c$, so wäre $y \in [< b] \cap B$ (wegen $c < b$) im Widerspruch zur Bedingung (+). □

Im folgenden schreiben wir $R$ für die Wohlordnung $<$ (denn wir benötigen nur die Fundiertheit von $R$):

### 6.1.3 Induktionsprinzip für Wohlordnungen

*$R$ sei Wohlordnung auf $A$. Dann gilt:*

(i) $\forall x \in A\, [\forall y(yRx \rightarrow \varphi(y)) \rightarrow \varphi(x)] \rightarrow \forall x \in A\, \varphi(x)$, *bzw. mit* $B = \{x \mid \varphi(x)\}$:

(ii) $\forall x \in A\, [S(x,R) \subseteq B \rightarrow x \in B] \rightarrow A \subseteq B.$

Dieses Prinzip bedeutet, daß man eine Aussage $\varphi(x)$ für alle Elemente von $x \in A$ nachweisen kann, wenn man für ein beliebiges $a \in A$ zeigen kann:

| | | |
|---|---|---|
| Falls | $\forall y(yRa \rightarrow \varphi(y))$ | **(Induktionsannahme)**, |
| so gilt | $\varphi(a)$ | **(Induktionsschluß)**. |

Dieses Beweisverfahren nennt man auch **Beweis durch $R$-Induktion**

*Beweis* von (ii): Annahme: $\forall x \in A\, [S(x,R) \subseteq B \rightarrow x \in B]$, aber $A \not\subseteq B$.

Dann ist $A - B$ nicht leer und besitzt nach obigem Satz ein $R$-minimales Element $a$:

$a \in A - B$, aber $\forall y \in A - B\, \neg yRa$, d. h. $\forall y(yRa \rightarrow y \in B)$,
also $S(a,R) \subseteq B$ und damit nach der Voraussetzung der $R$-Induktion: $a \in B$, Widerspruch!
(Tatsächlich ist die Aussage (i) von 6.1.2 logisch äquivalent zur Aussage (ii) dieses Satzes: benutze Kontraposition, logische Umformungen und gehe von $B$ zur Komplementklasse über!) □

### 6.1.4 Segmente

Für eine Relation $R$ definieren wir:

$R$-*Segment*$(B) :\leftrightarrow \forall x \forall y(xRy \wedge y \in B \rightarrow x \in B)$ **$R$-Segment**

Im Falle der $\in$-Beziehung $E$ gilt:

$$E\text{-}Segment(B) \leftrightarrow trans(B).$$

Für Ordnungen $<$ sind wegen der Transitivität die Mengen $S(a,<) = \{x \mid x < a\}$ offenbar Segmente (genannt *Anfangssegmente*), für Wohlordnungen sind es die einzigen echten Segmente:

#### Lemma

(i) *$R$ sei lineare Ordnung auf $A$, und $S,T \subseteq A$ seien Segmente. Dann gilt: $S \subseteq T \vee T \subseteq S$.*

(ii) *Ist $R$ Wohlordnung auf $A$, $S \subseteq A$ ein echtes Segment, so $S = S(a,R)$ für ein $a \in A$.*

*Beweis* von (i): Sei $S \subseteq T$ falsch. Dann gibt es ein $a \in S - T$, und dieses kann kein $R$-größeres $b \in T$ besitzen (sonst wäre auch $a \in T$, da $T$ ein Segment ist). Also gilt für alle $b \in T : bRa$, insbesondere $T \subseteq S$.

Für den Beweis von (ii) wähle man das $R$-kleinste Element $a \in A - S$. □

### 6.1.5 Rekursionssatz für Wohlordnungen

*Es sei $R$ eine Wohlordnung auf $A$, $G : V \times V \longrightarrow V$ eine Funktion. Dann existiert genau eine Funktion*

*$F : A \longrightarrow V$, die durch die folgenden rekursiven Eigenschaften bestimmt ist:*

$$F(a) = G(a, F \restriction S(a,R)).$$

Eine einfache Variante dieses Satzes liefert eine Funktion $F$ mit

$$F(a) = G(a, \{F(x) \mid xRa\}).$$

Ferner können $G$ und $F$ weitere Stellen als Parameter enthalten, z. B.

$$F(a,u,v) = G(a,u,v,\{F(x,u,v) \mid xRa\}).$$

*Beweis:*

a) Die Eindeutigkeit von $F$ zeigt man durch $R$-Induktion,

b) für den Nachweis der Existenz definiert man:

(i) h *brav:* $\leftrightarrow \exists s\,(s$ ist Segment $\wedge\, h : s \to V \wedge \forall x \in s\; h(x) = G(x, h \restriction S(x,R)))$,

(ii) *h verträglich* mit $g :\leftrightarrow \forall x \in D(h) \cap D(g)\; h(x) = g(x)$

Durch $R$-Induktion zeigt man, daß je zwei brave Funktionen miteinander verträglich sind:

(iii) $h, g$ *brav* $\to \forall x \in D(h) \cap D(g)\; h(x) = g(x)$.
(Mit Hilfe von (i) des vorigen Lemmas kann man sogar zeigen, daß für je zwei brave $h, g$ gilt $\quad g \subseteq h \vee h \subseteq g$.)

Die gesuchte Funktion können wir jetzt explizit definieren durch:

(iv) $F := \bigcup\{h \mid h \textit{ brav}\}$.

$F$ ist wegen (iii) eine Funktion, $D(F)$ ist ein Segment, und $F$ erfüllt die Rekursionsbedingungen (da Vereinigung braver Funktionen). Es bleibt zu zeigen, daß $F$ auf ganz $A$ definiert ist:

Wäre $B := D(F)$ echtes Segment von $A$, so $B = S(a,R)$ für ein $a \in A$ nach obigem Lemma (ii), insbesondere $B$ und damit $F$ eine Menge, etwa $F = f$ mit bravem $f$. Dieses $f$ könnte man dann fortsetzen zu einem braven
$h := f \cup \{(a, G(a, f \restriction S(a,R)))\}$
mit $a \in D(h) = S(a,R) \cup \{a\}$ ein Segment, also müßte auch $a \in D(F) = B = S(a,R)$ sein und somit $aRa$, Widerspruch! $\square$

Damit haben wir eine *explizite* mengentheoretische Definition einer Funktion erhalten, die durch eine rekursive Bedingung charakterisiert ist, und zwar erhält man die Funktion als Vereinigung von "partiellen" Lösungen der Rekursionsgleichungen. Insbesondere kann man damit die Elemente einer wohlgeordneten Klasse der Größe nach aufzählen - dabei hängt nämlich die Aufzählung eines Elementes $b \in A$ ab von der Aufzählung der Vorgänger von $b$:

**Kontraktionslemma**

*$R$ sei eine Wohlordnung auf $A$. Dann existiert genau eine Abbildung $F$ und eine transitive Klasse $B$, so daß:*

$$F \textit{ ein Isomorphismus: } (A,R) \cong (B,\in) \textit{ ist, d. h.}$$
$$F : A \longleftrightarrow B \quad \textit{mit}$$
$$aRb \leftrightarrow F(a) \in F(b) \quad \textit{für alle} \quad a,b \in A.$$

Dabei sei das **geordnete Paar von Klassen** etwa definiert durch:

$$(A,B) := (A \times \{0\}) \cup (B \times \{1\}),$$

und wir schreiben einfach

$(B,\in)$ für die Struktur $(B,E_B)$ mit $E_B := \{x,y \mid x,y \in B \wedge x \in y\}$.

*Beweis:* a) *Eindeutigkeit:*
Sei $F : A \longleftrightarrow B$ mit $aRb \leftrightarrow F(a) \in F(b)$, $B$ transitiv. Dann gilt für alle $b \in A$:

| | | |
|---|---|---|
| | $F(b) \in B \rightarrow F(b) \subseteq B$ | wegen trans(B) |
| | $a \in F(b) \rightarrow \exists x(a = F(x) \in F(b))$ | nach Voraussetzung |
| | $a \in F(b) \leftrightarrow \exists x(a = F(x) \wedge xRb)$, | somit |
| (*) | $F(b) = \{F(x) \mid xRb\}$. | |

Dadurch ist aber $F$ eindeutig festgelegt, und $B$ ist eindeutig als Wertebereich von F bestimmt.

b) *Existenz:* Nach dem Rekursionssatz existiert (genau) ein $F$ mit

$D(F) = A$ und für alle $b \in A$: $F(b) = \{F(x) \mid xRb\}$,

d. h. es existiert ein

$F : A \twoheadrightarrow B$ mit der Eigenschaft (*), wobei wir $B := W(F)$ gesetzt haben.

$F$ ist *injektiv:* Sei $F(y) = F(z)$, aber $y \neq z$. Da $R$ eine Ordnung ist, so gilt

$yRz$ und damit $F(y) \in F(z)$ oder

$zRy$ und damit $F(z) \in F(y)$,

in beiden Fällen ein Widerspruch zu $F(y) = F(z)$ (und dem Fundierungsaxiom)!

$F$ ist *Isomorphismus:* Nach Def. von $F$ gilt: $aRb \rightarrow F(a) \in F(b)$.
Sei umgekehrt $F(a) \in F(b)$. Dann ist $F(a) = F(c)$ für ein $cRb$, aber $a = c$ (wegen der Injektivität), und somit auch $aRb$.

$B$ ist *transitiv:* Sei $c \in B = W(F)$. Dann ist $c = F(b)$ für ein $b \in A$, andererseits nach Def. von $F$ (d.h. nach (*)): $c = F(b) = \{F(x) \mid xRb\} \subseteq B$, also $c \subseteq B$. □

Da transitive Mengen, die durch die $\in$-Beziehung wohlgeordnet sind, gerade die Ordinalzahlen sind, erhalten wir mit

$On := \{x \mid Ord(x)\}$ als **Klasse aller Ordinalzahlen**:

### 6.1.6 Repräsentationssatz für Wohlordnungen

*$<$ sei eine Wohlordnung auf $A$.*

(i) *Ist $A = a$ eine Menge, so gibt es genau eine Abbildung $f$ und genau eine Ordinalzahl $\alpha$ mit*

$$f : (a, <) \cong (\alpha, \in).$$

(ii) *Ist $A$ eine echte Klasse, so gibt es genau eine Abbildung $F$ mit*

$$F : (A, <) \cong (On, \in).$$

□

Somit sind die Ordinalzahlen Repräsentanten von Wohlordnungen. Das eindeutig bestimmte $\alpha$ mit $(a, <) \cong (a, \in)$ nennt man den **Wohlordnungstyp** von $(a, <)$. Man erhält ihn, wenn man die Elemente der Menge $a$ gemäß der Wohlordnung $<$ aufzählt: Ist

$$a = \{a_\xi \mid \xi < \alpha\} \text{ mit } a_0 < a_1 < \ldots < a_\xi < a_\eta \ldots \text{ für } \xi < \eta < \alpha,$$

so ist $\alpha$ der Wohlordnungstyp von $a$ bezüglich der Wohlordnung $<$ auf $a$.

Im Falle einer echten Klasse gibt es nur einen Wohlordnungstyp (nämlich $On$), und zwar liegt das an der Mengenbedingung (F1) in der Definition 6.1.1.3.

#### Anwendung auf die Ordinalzahlen

Die kleinste Ordinalzahl ist 0, und für jede Ordinalzahl $\alpha$ ist $\alpha + 1$ ebenfalls eine Ordinalzahl, und zwar eine Nachfolgerzahl. Die restlichen Typen von Ordinalzahlen sind Limeszahlen:

$$\begin{aligned} 0 &:= \emptyset && \textbf{Null} \\ Nf(\alpha) &: \leftrightarrow \exists \xi\ \alpha = \xi + 1 && \textbf{Nachfolgerzahl} \\ Lim(\lambda) &: \leftrightarrow \lambda \neq 0 \wedge \neg Nf(\lambda)) && \textbf{Limeszahl} \\ &\leftrightarrow \lambda \neq 0 \wedge \forall \xi < \lambda\ (\xi + 1 < \lambda) \end{aligned}$$

#### Lemma

(i) $\alpha = 0 \vee Nf(\alpha) \vee Lim(\alpha)$ *(und es gilt genau einer der drei Fälle),*

(ii) $trans(A) \wedge A \subseteq On \to A \in On \vee A = On,$

(iii) $A \subseteq On \rightarrow \cup A \in On \vee \cup A = On$.

(iv) $Lim(\lambda) \leftrightarrow \lambda > 0 \wedge \lambda = \cup\lambda$.

*Beweis* von (ii): Ist $A$ transitiv, so ein E-Segment von $On$ und nach Lemma (ii) aus 6.1.4 somit $= On$ oder als echtes Segment von der Form $A = \{x \mid x < \alpha\} = \alpha$ für ein $\alpha$.
Zum Beweis von (iii) beachte man, daß für $A \subseteq On$ gilt: $\cup A$ ist transitiv (allgemeiner ist die Vereinigung transitiver Mengen wieder transitiv), so daß man (ii) anwenden kann. □

Offenbar ist (ii) eine Verallgemeinerung des Satzes aus 5.1.4. (iii) besagt, daß das Supremum einer Klasse von Ordinalzahlen wieder eine Ordinalzahl ist oder = On ist, und zwar gilt in (ii) und (iii) jeweils der 1. Fall, wenn $A$ beschränkt ist (d. h. $\exists\alpha\, A \subseteq \alpha$), der 2. Fall, falls $A$ unbeschränkt ist (also $\forall\alpha\exists\xi \geq \alpha\ \xi \in A$).

Die Existenz von Limeszahlen folgt erst aus dem Unendlichkeitsaxiom, und damit werden wir uns im folgenden Abschnitt beschäftigen. Hier können wir die Existenzfrage erst einmal offen lassen; falls es keine Limeszahl gibt, sind die Ordinalzahlen gerade die natürlichen Zahlen, anderenfalls gehen sie darüber hinaus: Mit einer Limeszahl $\lambda$ gibt es auch wieder die Nachfolgerzahlen $\lambda+1, \lambda+2, \ldots$ (die aber keine natürlichen Zahlen mehr sind) und dann wieder deren Limes, Nachfolger, Nachfolger, ..., Limes, ..., usw. Somit führen die Ordinalzahlen über die natürlichen Zahlen hinaus, indem zu jeder Menge von Ordinalzahlen auch das Supremum wieder eine Ordinalzahl ist: Es gilt

$$(*) \quad 0 = \emptyset \in On \wedge \forall\alpha(\alpha \in On \rightarrow \alpha+1 \in On) \wedge \forall x(x \subseteq On \rightarrow \cup x \in On),$$

tatsächlich ist $On$ die kleinste Klasse mit dieser Abgeschlossenheitsbedingung.

Da die Ordinalzahlen durch die $\in$ -Beziehung wohlgeordnet sind, gelten die Induktions- und Rekursionstheoreme des vorigen Abschnitts; zusätzlich kann man ihnen aber wegen (*) auch eine zweite Fassung geben, die zeigt, wie diese Prinzipien die entsprechenden Sätze über die natürlichen Zahlen verallgemeinern:

### Prinzip der transfiniten Induktion/Minumumsprinzip

(i) $\forall\alpha(\alpha \subseteq A \rightarrow \alpha \in A) \rightarrow On \subseteq A$

(ii) $\varphi(0) \wedge \forall\alpha((\varphi(\alpha) \rightarrow \varphi(\alpha+1)) \wedge$
$\wedge \forall\lambda(Lim(\lambda) \wedge \forall\xi < \lambda\, \varphi(\xi) \rightarrow \varphi(\lambda)) \rightarrow \forall\alpha\varphi(\alpha)$

(iii) $\exists\alpha\varphi(\alpha) \rightarrow \exists\alpha(\varphi(\alpha) \wedge \forall\xi < \alpha\neg\varphi(\xi))$ □

Dabei haben wir in (i) die Klassen-, in (ii) die Schema-Schreibweise benutzt. (iii) ist das hierzu äquivalente Minimumsprinzip. Im folgenden Satz beschränken wir uns auf den Fall, welcher der Induktion der Form (ii) entspricht:

### 6.1.7 Transfinite Rekursion

*Sind $G,H : On \times V \longrightarrow V$ Funktionen, a eine Menge, so existiert genau eine Funktion $F : On \longrightarrow V$ mit*

$$\begin{aligned} F(0) &= a \\ F(\alpha+1) &= G(\alpha, F(\alpha)) \\ F(\lambda) &= H(\lambda, \{F(\xi) \mid \xi < \lambda\}) \quad \text{für } Lim(\lambda). \end{aligned}$$

(Analog für mehrstellige Funktionen mit Parametern.) □

Der Rekursionssatz erlaubt es, die arithmetischen Operationen auf den Ordinalzahlen so einzuführen, daß sie die entsprechenden Operationen auf den natürlichen Zahlen verallgemeinern (nach JACOBSTHAL 1909):

**Addition, Multiplikation, Potenz**

$$\begin{array}{ll} \alpha+0=\alpha & \alpha\cdot 0=0 \\ \alpha+(\beta+1)=(\alpha+\beta)+1 & \alpha\cdot(\beta+1)=\alpha\cdot\beta+\alpha \\ \alpha+\lambda=\bigcup_{\xi<\lambda}\alpha+\xi, \ \text{falls Lim}(\lambda) & \alpha\cdot\lambda=\bigcup_{\xi<\lambda}\alpha\cdot\xi, \ \text{falls Lim}(\lambda) \end{array}$$

$$\begin{aligned} \alpha^0 &= 1 \\ \alpha^{(\beta+1)} &= \alpha^\beta \cdot \alpha \\ \alpha^\lambda &= \textstyle\bigcup_{0<\xi<\lambda} \alpha^\xi, \quad \text{falls Lim}(\lambda) \end{aligned}$$

Die von den natürlichen Zahlen bekannten Rechengesetze gelten zum Teil auch für die unendlichen Ordinalzahlen ($+$ und $\cdot$ sind assoziativ mit 0 bzw. 1 als neutralen Elementen, alle Operationen sind schwach monoton), Addition und Multiplikation sind jedoch nicht kommutativ und verhalten sich nur hinsichtlich ihres 2. Argumentes wie gewohnt. Soweit die üblichen Rechengesetze jedoch nicht gelten, sind sie bereits im kleinst-möglichen Fall verletzt, z.B.:

**nicht-kommutativ:** $1+\omega=\omega\neq\omega+1,\ 2\cdot\omega=\omega\neq\omega\cdot 2=\omega+\omega$

**nicht-links-distributiv:** $(1+1)\cdot\omega=\omega\neq 1\cdot\omega+1\cdot\omega=\omega+\omega$

**nicht-monoton im 1. Argument:** $1+\omega=2+\omega,\ 1\cdot\omega=2\cdot\omega,\ 2^\omega=3^\omega$

## 6.2 Die von-Neumannsche Hierarchie

Durch transfinite Rekursion definieren wir eine Folge von Mengen $(V_\alpha | \alpha \in On)$, indem wir ausgehend von der leeren Menge die Potenzmengenoperation iterieren:

$$V_0 = \emptyset,\ V_{\alpha+1} = \mathcal{P}(V_\alpha),\ V_\lambda = \bigcup_{\xi<\lambda} V_\xi \quad \text{für Limeszahlen } \lambda.$$

Die Mengen $V_\alpha$ nennt man auch VON NEUMANNsche **Stufen**. Das Fundierungsaxiom ist äquivalent damit, daß jede Menge als Element einer Stufe vorkommt, d. h. mit der Aussage

$$V = \bigcup_{\alpha \in On} V_\alpha.$$

Damit erhält man eine Einteilung der Klasse $V$ aller Mengen in eine Hierarchie von Stufen, die gewisse "Ähnlichkeiten" mit $V$ aufweisen. Genauer läßt sich dieser Zusammenhang in der Form eines

**Reflexionsprinzip** (Refl)
$$\exists \alpha [a \in V_\alpha \wedge \forall x_1 \dots x_n \in V_\alpha (\varphi^{V_\alpha}(x_1 \dots x_n) \leftrightarrow \varphi(x_1 \dots x_n))]$$

ausdrücken, wobei die Formel $\varphi^a$ (die *Relativierung* von $\varphi$ nach $a$) aus $\varphi$ hervorgeht, indem dort jeder Quantor $Qx$ durch den relativierten Quantor $Qx \in a$ ersetzt wird. Somit besagt das Reflexionsprinzip, daß zu jeder Formel $\varphi$ ein beliebig großes $V_\alpha$ existiert, das bezüglich der Eigenschaft $\varphi$ sich wie die Allklasse $V$ verhält. Man kann Reflexionsprinzipien benutzen als alternative (und weniger pragmatische) Möglichkeit, die Theorie ZF zu axiomatisieren - bei geeigneter Formulierung machen sie alle Mengenexistenz-Axiome (bis auf das Aussonderungsaxiom) überflüssig und sind natürlich modelltheoretisch von besonderem Interesse.

Die Stufen $V_\alpha$ sind transitive Mengen mit interessanten Abschlußeigenschaften:

- $(V_\omega, \in)$ ist Modell aller Axiome von ZF, bis auf das Unendlichkeitsaxiom, das hierin falsch ist (da alle Mengen in $V_\omega$ endlich sind, s. unten 6.3.2).
- $(V_\lambda, \in)$ ist für Limeszahlen $\lambda > \omega$ Modell aller Axiome von ZF, bis auf das Ersetzungsaxiom.
- Die Frage nach der Existenz von Ordinalzahlen $\alpha$, für die $(V_\alpha, \in)$ ein Modell von *ZF* ist, führt zu den "großen Kardinalzahlen", deren Existenz in ZF nicht beweisbar ist ("unerreichbare" Zahlen).

## 6.3 Die Rolle des Unendlichkeitsaxioms

Wenn man möglichst schnell und einfach das Induktionsprinzip für die natürlichen Zahlen erhalten will, so definiert man sie als Durchschnitt aller induktiven Mengen (und damit als kleinste induktive Menge):

$$\mathbb{N} := \bigcap\{x \mid Ind(x)\}, \quad \text{wobei}$$
$$Induktiv(a) :\leftrightarrow \emptyset \in a \wedge \forall y(y \in a \rightarrow y \cup \{y\} \in a)).$$

Da der Durchschnitt über die leere Menge jedoch keine Menge (sondern die echte Klasse aller Mengen) ist, muß man für diesen Weg das Unendlichkeitsaxiom voraussetzen (welches gerade besagt, daß es eine induktive Menge gibt). Wir wollen jedoch - soweit möglich - dieses Axiom vermeiden und wählen daher den *Begriff* der natürlichen Zahl als "endliche" Ordinalzahl, d. h. als Ordinalzahl unter der ersten Limeszahl:

$$\begin{aligned} Nz(n): \quad &\leftrightarrow \quad Ord(n) \wedge (n = 0 \vee Nf(n)) \wedge \forall x \in n\,(x = 0 \vee Nf(x)) \\ &\leftrightarrow \quad Ord(n) \wedge \forall \lambda (Lim(\lambda) \rightarrow n < \lambda) \quad \textbf{natürliche Zahl} \end{aligned}$$

Da das Unendlichkeitsaxiom (Un) auch damit gleichbedeutend ist, daß eine Limeszahl existiert, bedeutet diese Definition:

a) Gilt (Un), so ist die kleinste Limeszahl $\omega =$ *Menge* der natürlichen Zahlen.

b) Gilt $\neg$(Un), existiert also keine unendliche Menge, so sind alle Mengen endlich und die natürlichen Zahlen bilden eine echte *Klasse*, die mit der Klasse *On* aller Ordinalzahlen zusammenfällt.

Da somit die natürlichen Zahlen spezielle Ordinalzahlen sind, ergeben sich aus den jeweiligen Prinzipien für die Ordinalzahlen entsprechende Aussagen über die natürlichen Zahlen (dabei benutzen wir wie üblich Variable n,m,k für natürliche Zahlen):

**Induktionsprinzip** $\forall n(\forall m < n\; m \in A \rightarrow n \in A) \rightarrow \forall n\; n \in A$

**Induktionsschema** $\varphi(0) \wedge \forall n((\varphi(n) \rightarrow \varphi(n+1)) \rightarrow \forall n\; \varphi(n)$

**Minimumsprinzip** $\exists n\, \varphi(n) \rightarrow \exists n(\varphi(n) \wedge \forall m < n \neg \varphi(m))$

**Rekursion** *Sind $G,H:\mathbb{N}\times V\longrightarrow V$ Funktionen, $a$ eine Menge, so existiert genau eine Funktion $F:\mathbb{N}\longrightarrow V$ mit*

$$\begin{aligned} F(0) &= a \\ F(n+1) &= G(n,F(n)). \end{aligned}$$

### 6.3.1 PA in mengentheoretischer Sprache

Die Theorie PA der Zahlentheorie von PEANO haben wir bereits im Rahmen der Logik (3.5.4) erwähnt und festgestellt, daß diese Theorie neben dem Standardmodell weitere *Nichtstandardmodelle* besitzt (als Folgerung aus dem Kompaktheitssatz). Der Versuch der Erweiterung von PA zur Theorie $\mathsf{PA}^{(2)}$ in der Sprache der 2. Stufe trifft auf die Schwierigkeit, daß die Logik 2. Stufe kein vollständiges Axiomensystem besitzt (zumindest solange man verlangt, daß die Menge der Axiome entscheidbar ist). Gehen wir nun jedoch von der Mengenlehre aus, so erhalten wir zunächst die Möglichkeit, das Standardmodell zu definieren als Menge der natürlichen Zahlen $\omega$ mit $a' = a\cup\{a\}$ , $0=\emptyset$ und mit den mengentheoretisch (durch Rekursion auf den natürlichen Zahlen) definierten Operationen $+,\cdot$ (der Einfachheit halber benutzen wir hierfür dieselben Zeichen wie für die Symbole der Sprache von PA):

**Standardmodell** $\quad\mathcal{N} = (\omega,',+,\cdot,0).$

Damit dieses eine Menge ist, müssen wir natürlich das Unendlichkeitsaxiom voraussetzen; dann aber gilt das Induktionsschema im Standardmodell nicht nur für Aussagen der Sprache von PA, sondern für beliebige *mengentheoretische* Formeln, und es läßt sich in ZF beweisen, daß das Standardmodell bis auf Isomorphie das einzige dieser (mengentheoretisch erweiterten) Axiome ist (wobei man P3-P6 nicht einmal benötigt):

#### Kategorizität der mengentheoretischen Peano-Axiome

Die Theorie $\mathsf{PA}^{(2)}$ (PEANO-Axiome 2. Stufe) besitzt folgende Axiome:

P1 $\quad\forall x\, e\neq x$

P2 $\quad\forall x\forall y(x'=y'\rightarrow x=y)$

$\mathrm{Ind}^{(2)}$ $\quad\forall X(0\in X\wedge\forall x(x\in X\rightarrow x'\in X)\rightarrow\forall x\, x\in X)$

Die *Kategorizität* dieser Theorie besagt, daß es (bis auf Isomorphie) genau ein Modell gibt, d. h.:

*Jedes Modell von* $\mathsf{PA}^{(2)}$ *ist isomorph zum Standardmodell* $(\omega,',0)$.

*Beweis:* $\mathsf{Ind}^{(2)}$ ist ein Axiom, in welchem $\forall X$ ein Quantor 2. Stufe ist, welcher in einem Modell $(w,*,e)$ ausdrückt: *für alle Teilmengen von w ...*, d. h. dieses Axiom entspricht dem mengentheoretische Induktionsprinzip.

Es sei nun also $(w,^*,e)$ ein Modell von $\mathsf{PA}^{(2)}$, d.h.

(1) $e \neq x^*$ für alle $x \in w$,

(2) $x^* = y^* \rightarrow x = y$ für alle $x, y \in w$,

(3) $e \in v \wedge \forall x(x \in v \rightarrow x^* \in v) \rightarrow w \subseteq v$ für alle Mengen $v \subseteq w$.

Der gesuchte Isomorphismus ist eine Abbildung $f : \omega \longleftrightarrow w$ mit

$$f(0) = e \wedge \forall n \in \omega\; f(n') = f(n)^*.$$

Der Rekursionssatz für die natürlichen Zahlen liefert ein (sogar eindeutig bestimmtes) $f$ mit dieser Eigenschaft; zu zeigen bleibt also nur noch die Bijektivität:

$f$ ist surjektiv: Sei $v := W(f) = \{f(n) \mid n \in \omega\}$. Dann folgt aus den rekursiven Bedingungen für $f$:

$$e \in v \wedge \forall x(x \in v \rightarrow x^* \in v),$$

also nach (3): $w \subseteq v$ und damit $w = v$, d.h. $f$ ist surjektiv.

$f$ ist injektiv, d.h.

$$(i) \quad m \neq n \rightarrow f(m) \neq f(n).$$

Zunächst zeige man durch Induktion über m:

$$(ii) \quad m \neq 0 \rightarrow \exists n(m = n')$$

und dann (i) durch Induktion über n:

1. Sei $n = 0$ und $n \neq m$. Dann ist $m \neq 0$, also nach (i) $m = k'$ für ein $k$ und damit $f(0) = e \neq f(m) = f(k)^*$ nach (2).

2. Sei (i) bewiesen für $n$ (und alle $m$), $m \neq n'$. Falls $m = 0$, schließen wir wie oben auf $f(m) \neq f(n')$. Falls $m \neq 0$, so wieder $m = k'$ für ein $k$. Aus $m = k' \neq n'$ folgt wegen (P2) $k \neq n$ und damit nach Induktionsvoraussetzung $f(k) \neq f(n)$.

$\square$

Hier zeigt sich besonders deutlich, daß die zahlentheoretische Sprache (1. Stufe) erheblich ausdrucksschwächer ist als die mengentheoretische Sprache.

### 6.3.2 Die Theorie der endlichen Mengen

Bezeichnet $\mathsf{ZF}_{fin}$ die Theorie ZF, in welcher das Unendlichkeitsaxiom durch seine Negation ersetzt ist, so kann man sie als *Theorie der endlichen Mengen* auffassen; $(V_\omega, \in)$ ist "das Standardmodell" dieser Theorie. Diese ist ebenso "stark" wie die PEANO-Arithmetik PA, d. h. sie lassen sich wechselweise ineinander interpretieren, und damit ist insbesondere ZF ohne das Unendlichkeitsaxiom genau dann widerspruchsfrei, wenn die PEANO-Arithmetik widerspruchsfrei ist:

- Wie oben bemerkt, kann man in ZF *ohne Unendlichkeitsaxiom* die natürlichen Zahlen (als Klasse) definieren und die arithmetischen Operationen der Addition und Multiplikation auf den natürlichen Zahlen (mittels des Rekursionssatzes) definieren, so daß hierfür die PEANO-Axiome gelten,
- umgekehrt kann man in PA (nach ACKERMANN) endliche Mengen durch Zahlen kodieren und eine entsprechende $\in$-Beziehung so definieren, daß hierfür die Axiome von $\mathsf{ZF}_{fin}$ gelten.
- In ZF *ohne Unendlichkeitsaxiom* (oder einer geeigneten Erweiterung von PA zur Theorie 2. Stufe) kann man auch eine eingeschränkte Analysis betreiben; um aber etwa die Gesamtheit der reellen Zahlen zur Verfügung zu haben, benötigt man das Unendlichkeitsaxiom.

### 6.3.3 Anwendungen der numerischen Rekursion

Als Spezialfall des Rekursionstheorems für die natürlichen Zahlen erhalten wir die Möglichkeit, Funktionen zu iterieren und damit Mengen zu erhalten, die unter vorgegebenen Funktionen abgeschlossen sind:

**Satz über die Iteration und den Abschluß**

(i) *Zu jeder Funktion F und jeder Menge a gibt es eine Funktion*

$$H : \mathbb{N} \times V \longrightarrow V \text{ mit } H(0,a) = a,\ H(n+1,a) = F(H(n,a)).$$

Wir nennen $H$ auch die **Iteration von** $G$ **mit dem Anfangswert** $a$ und schreiben auch $F^n(a)$ für $H(n,a)$.

(ii) *Zu jeder Funktion F und jeder Menge a gibt es eine Obermenge b von a, die unter F abgeschlossen ist:*

$$\exists y(a \subseteq y \wedge \forall x \in y\, F(x) \in y)$$

*Beweis* von (ii): Zu gegebenem $F$ und der Ausgangsmenge $a$ definiere durch Rekursion eine Funktion $G$ mit den Eigenschaften

$$G(0) = a, \quad G(n') = \{F(x) \mid x \in G(n)\} \quad \text{und setze} \quad b := \bigcup_{n \in \omega} G(n).$$

Dann ist offensichtlich $a \subseteq b \wedge \forall x \in b\, F(x) \in b)$, und zwar ist $b$ die kleinste derartige Menge, die **Hülle** oder der **Abschluß** von $a$ unter der Abbildung $F$. □

Für viele Anwendungen gut zu benutzen ist das

**Hüllenaxiom** (HS) $\quad \exists y(a \subseteq y \wedge \forall x \in y\, F(x) \in y)$.

**Satz**

*Unter der Voraussetzung der übrigen Axiome von* ZF *(einschließlich des Aussonderungsschemas) ist das*

Ersetzungsaxiom + Unendlichkeitsaxiom *äquivalent zum* Hüllenaxiom.

*Beweis:* (HS) haben wir in obigem Satz in ZF gezeigt. Umgekehrt folgt aus (HS) das Unendlichkeitsaxiom (mit $a = \emptyset$ und $F(x) = x'$) und ähnlich das Ersetzungsaxiom, indem es für eine Funktion $F$ und eine Menge $a$ zunächst eine Obermenge $b \supseteq a$ mit $\{F(x)|x \in b\} \subseteq b$ liefert, für die dann aber $\{F(x) \mid x \in a\}$ als Teilklasse von $b$ eine Menge nach dem Aussonderungsaxiom ist. □

## Anwendungen des Hüllenaxioms

1. Ähnlich wie wir $\omega$ als (kleinste) Ordinalzahl erhalten können, die $> 0$ ist und unter $'$ abgeschlossen ist, erhalten wir zu jeder Ordinahlzahl $\alpha$ eine (nächstgrößere) Limeszahl $\lambda > \alpha$, aber auch Ordinalzahlen $> \alpha$, die unter arithmetischen Operationen abgeschlossen sind:
$\forall \xi, \eta < \gamma \quad \xi + \eta < \gamma$ bzw. die **$\varepsilon$-Zahlen:** $\forall \xi < \varepsilon \; 2^{\xi} < \varepsilon$, usw.

2. Zu jeder Menge $a$ existiert eine kleinste Obermenge $b \supseteq a$ mit $trans(b)$, die **transitive Hülle** $TC(a)$ von a:

$$TC(a) := \bigcup_{n \in \omega} U^n(a), \quad \text{wobei } U \text{ die Iteration der}$$

$$\text{Vereinigungsmenge ist: } U^o(a) = a, \; U^{n+1}(a) = \cup U^n(a),$$

$$\text{also} \quad TC(a) = a \cup \bigcup a \cup \bigcup \bigcup a \ldots$$

3. Eine Menge ist **endlich** gdw sie sich auf die Zahlen $\{0, \ldots, n\}$ für eine natürliche Zahl $n$ abbilden läßt,

$$HF := \{x \mid TC(x) \; endlich\} \quad \text{ist die Klasse der } \textbf{erblich-endlichen} \text{ Mengen.}$$

Sie läßt sich auch beschreiben durch $HF = V_\omega$ und ist somit eine Menge (wobei man wesentlich das Unendlichkeitsaxiom benutzt - dagegen ist die Klasse aller endlichen Mengen stets eine echte Klasse).

4. Mit Hilfe der VON NEUMANNschen Stufenhierarchie kann man die folgenden Verstärkungen des Hüllenaxioms erhalten

$$\forall x \exists y \, \varphi(x, y, \ldots) \rightarrow \exists \alpha \forall x \in V_\alpha \exists y \in V_\alpha \; \varphi^{V_\alpha}(x, y, \ldots), \text{ bzw.}$$
$$\forall x \in a \exists y \, \varphi(x, y, \ldots) \rightarrow \exists u (a \subseteq u \land \forall x \in a \exists y \in u \; \varphi^u(x, y, \ldots)),$$

welche sich auch zur Axiomatisierung von ZF verwenden lassen (vgl. die in 6.2 erwähnten Reflexionsprinzipien).

# Kapitel 7

# Das Auswahlaxiom

## 7.1 Zermelos Axiom

In seiner einfachsten Form (RUSSELL 1906 bzw. ZERMELO 1904/1908) besagt das Auswahlaxiom, daß das Produkt einer Menge nicht-leerer Mengen wiederum nicht leer ist (*multiplicative axiom*):

**Auswahlaxiom:** (AC) $\forall x \in a\, x \neq \emptyset \to \prod_{x \in a} x \neq \emptyset$,

d. h. $\forall x \in a\, x \neq \emptyset \to \exists f (Fkt(f) \wedge D(f) = a \wedge \forall x \in a\, f(x) \in x)$,

(ein solches $f$ heißt **Auswahlfunktion** für die Menge $a$).

Das AC ist unabhängig von den übrigen $\mathsf{ZF}$-Axiomen:
Ist $\mathsf{ZF}_0 (= \mathsf{ZF}$ ohne Fundierungsaxiom) widerspruchsfrei, so auch

$\mathsf{ZF}_0 + \neg$ AC (FRAENKEL 1922),
$\mathsf{ZF}+$ AC (GÖDEL 1938),
$\mathsf{ZF} + \neg$ AC (COHEN 1963)

(zur Geschichte und Problematik des Auswahlaxioms s. das Buch von G. MOORE: *Zermelo´s axiom of choice.* Springer 1982).

Das Auswahlaxiom hat zahlreiche Anwendungen in fast allen Gebieten der Mathematik, besitzt aber auch verschiedene

### 7.1.1 Mengentheoretisch äquivalente Formen

AC2 $\forall x \in a\, F(x) \neq \emptyset \to \prod_{x \in a} F(x) \neq \emptyset$

AC2´ $\forall x \in a\, \exists y\, \varphi(x,y) \to \exists f (Fkt(f) \wedge D(f) = a \wedge \forall x \in a\, \varphi(x, f(x)))$

AC3 $\forall x \in a\, x \neq \emptyset \wedge \forall x, y \in a(x \neq y \rightarrow x \cap y = \emptyset) \rightarrow \exists z \forall x \in a \exists! u\, u \in z \cap x$
(ein solches $z$ heißt **Auswahlmenge** für die nicht-leeren, disjunkten Mengen in $a$; $\exists! u \ldots$ = es existiert genau ein $u \ldots$)

AC3´ $r$ Äquivalenzrelation auf $a \rightarrow \exists z \forall x \in a \exists! u \in z\,(x, u) \in r$
(jede Äquivalenzrelation besitzt ein Repräsentantensystem)

AC4 $Fkt(f) \rightarrow \exists g(Fkt(g) \wedge g : W(f) \rightarrowtail D(f) \wedge g \subseteq f^{-1})$
(jede Funktion besitzt eine Umkehrfunktion)

AC5 $Rel(r) \rightarrow \exists f(Fkt(f) \wedge D(f) = D(r) \wedge f \subseteq r)$

Der *Beweis* dieser Aussagen aus dem Auswahlaxiom beruht darauf, dass man in allen diesen Fällen eine Auswahl in Abhängigkeit von den Elementen einer *Menge* treffen muss (dabei benötigt man für AC2´ das Fundierungsaxiom, um den Bereich der möglichen $y$ auf eine genügend große Menge $V_\alpha$ einschränken zu können). Umgekehrt sind die obigen Auswahlprinzipien allgemein genug, um hieraus das ursprüngliche AC zu folgern. □

### 7.1.2 Der Zermelosche Wohlordnungssatz

*Das Auswahlaxiom,* AC, *ist äquivalent zum Wohlordnungssatz:*

WO1 $\forall x \exists r(r$ ist Wohlordnung auf $x)$, bzw.

WO2 $\forall x \exists f \exists \alpha(f : \alpha \longleftrightarrow x)$ , d. h. jede Menge $a$ läßt sich aufzählen:
$a = \{a_\xi \mid \xi < \alpha\}$ für ein $\alpha$ und eine Folge $(a_\xi \mid \xi < \alpha)$

*Beweis:* Ist $r$ Wohlordnung auf $a$, so isomorph zu $\in$ auf einer Ordinalzahl $\alpha$ (nach 6.1.6 (i)), somit WO1 $\Rightarrow$ WO2. Ist umgekehrt $f : a \longleftrightarrow \alpha$ für eine Ordinalzahl $\alpha$, so kann man auf $a$ eine Wohlordnung $r$ definieren durch

$$xry \iff f(x) < f(y).$$

Also gilt auch WO2 $\Rightarrow$ WO1.

WO1 $\Rightarrow$AC: (Ist eine Menge wohlgeordnet, so kann man aus ihren Elementen eine Auswahl treffen, indem man das kleinste Element (bezüglich der Wohlordnung) auswählt.) Sei $\forall x \in a\, x \neq \emptyset$. Nach WO1 gibt es eine Wohlordnung $<$ auf $\bigcup a$. Damit kann man eine Auswahlfunktion $f$ definieren durch

$$f(x) = \text{das kleinste (bzgl. } <) \; y \in x \text{ für } x \in a.$$

AC $\Rightarrow$ WO2: Sei $a \neq \emptyset$. Nach dem Auswahlaxiom existiert eine Funktion

$$f : \mathcal{P}(a) - \{\emptyset\} \longrightarrow V \text{ mit } \forall x \subseteq a (x \neq \emptyset \rightarrow f(x) \in x).$$

Damit können wir rekursiv eine Aufzählung $G$ von $a$ definieren:

$$G(\alpha) = \begin{cases} f(a - \{G(\xi) \mid \xi < \alpha\}), & \text{falls } a - \{G(\xi) \mid \xi < \alpha\} \neq \emptyset \\ a, & \text{sonst.} \end{cases}$$

(i) $a \notin \{G(\xi) \mid \xi < \alpha\} \rightarrow G \restriction \alpha$ ist injektiv:

Für $\gamma < \alpha$ ist $G(\gamma) = f(a - \{G(\xi) \mid \xi < \gamma\}) \in a - \{G(\xi) \mid \xi < \gamma\}$, also $G(\gamma) \neq G(\xi)$ für alle $\xi < \gamma$.

(ii) $a \in W(G)$, denn anderenfalls wäre $G : On \longrightarrow a$ nach (i) eine injektive Funktion, und damit $a$ keine Menge!

Sei $\alpha := \mu\xi(G(\xi) = a)$[1]. Dann ist $G \restriction \alpha : \alpha \longrightarrow a$ injektiv. Es bleibt zu zeigen:

(iii) $G \restriction \alpha : \alpha \longrightarrow a$ ist surjektiv. Wäre jedoch $W(G \restriction \alpha) \subset a$, so $G(a) = f(a - \{G(\xi) \mid \xi < \alpha\}) \in a$ im Widerspruch zur Definition von $\alpha$! $\square$

### 7.1.3 Maximumsprinzipien von Zorn und Hausdorff

Diese Prinzipien formulieren wir am besten für reflexive Ordnungen:

Eine (reflexive) **teilweise Ordnung** auf $A$ ist eine 2-stellige Relation $R$ auf $A$, für die gilt:

(a) $\forall x \in a \quad xRx$ **reflexiv**,
(b) $\forall x, y, z \in a\, (xRy \wedge yRz \rightarrow xRz)$ **transitiv**,
(c) $\forall x, y \in a\, (xRy \wedge yRx \rightarrow x = y)$ **antisymmetrisch**.
(d) $K \subseteq A$ heißt $R$-**Kette:** $\leftrightarrow \forall x, y \in K\, (xRy \vee yRx)$,
d.h. je zwei Elemente aus $K$ sind bzgl. $R$ vergleichbar;
(e) $a \in A$ heißt $R$-**obere Schranke** von $B \subseteq A :\leftrightarrow \forall x \in B\, xRa$,
$a \in A$ heißt $R$-**maximal:** $\leftrightarrow \forall x \in A\, (aRx \rightarrow a = x)$.

Ein maximales Element besitzt also kein echt größeres, braucht aber nicht das größte Element zu sein (zumindest nicht in einer teilweisen Ordnung).

**Zornsches Lemma** (ZL)

*r sei teilweise Ordnung auf der Menge a mit der Eigenschaft*
*(*) jede r-Kette besitzt eine r-obere Schranke.*
*Dann hat a ein maximales Element (bzgl. r).*

[1]Mit $\mu\xi\varphi(\xi)$ bezeichnet man das kleinste $\xi$ mit der Eigenschaft $\varphi(\xi)$.

**Hausdorffsches Maximumsprinzip** (H)

*r sei teilweise Ordnung auf der Menge a.*
*Dann gibt es eine (bzgl. $\subseteq$) maximale r-Kette k, d. h. ein $k \subseteq a$ mit:*
*k ist r-Kette $\wedge \forall y$ ( y r-Kette $\wedge k \subseteq y \to k = y$).*

Diese beiden Aussagen sind äquivalent zum Auswahlaxiom:

*Beweis* von WO $\Rightarrow$ H:
$r$ sei partielle Ordnung auf $a$, $a = \{a_\xi \mid \xi < \alpha\}$ sei wohlgeordnet. Wir definieren eine Teilmenge $\{b_\xi \mid \xi < \alpha\}$ von $a$ durch Rekursion wie folgt:

$$b_\beta = \begin{cases} a_\beta & \text{falls } \{b_\xi \mid \xi < \beta\} \cup \{a_\beta\} \text{ } r\text{-}Kette, \\ a_0 & \text{sonst.} \end{cases}$$

$\{b_\xi \mid \xi < \alpha\}$ ist dann eine maximale $r$-Kette (mit $b_0 = a_0$).

H $\Rightarrow$ ZL:
Sei $r$ partielle Ordnung auf $a$, die (*) erfüllt. Nach H existiert eine maximale $r$-Kette $k \subseteq a$. Eine obere Schranke von $k$ ist dann ein maximales Element (sonst könnte man $k$ echt erweitern).

ZL $\Rightarrow$ AC:
Sei $a \neq \emptyset$ und $\forall x \in a\, x \neq \emptyset$. Setze

$$B := \{f \mid \exists y \subseteq a (f : y \to \bigcup a \wedge \forall x \in y\, f(x) \in x)\}.$$

Jedes Element von $B$ ist eine partielle Auswahlfunktion und Teilmenge von $a \times \bigcup a$, damit ist auch $B$ als Menge $b \subseteq \mathcal{P}(a \times \bigcup a)$ abschätzbar. Als teilweise Ordnung auf $B = b$ wählen wir $r = \subseteq$ -Beziehung. Für jede $r$-Kette $k$ ist dann $\bigcup k$ eine obere Schranke, und ein nach dem ZORNschen Lemma maximales Element in $B$ ist dann eine Auswahlfunktion für $a$ . □

**Bemerkungen**

- Aus obigem Beweis lässt sich ablesen, daß H sogar allgemeiner gilt: jede partielle Ordnung $r$ auf $a$ besitzt für jedes $b \in a$ eine r-Kette k mit $b \in k$.
- Das ZORNsche Lemma ist offenbar trivial, wenn $r$ eine lineare Ordnung auf $a$ ist, weil dann $a$ selbst schon eine Kette ist. Daher kann man Anwendungen von ZL nur für teilweise Ordnungen erwarten. Wie in obigem Beweis ist häufig diese Ordnung die $\subseteq$ -Beziehung, und häufig kann man in diesem Fall (*) nachweisen, indem man zeigt, daß

$$\text{für jede } \subseteq\text{-Kette } k \text{ ist } \bigcup k \in a.$$

(Dann ist $\bigcup k$ nämlich eine obere Schranke.)

- Für Anwendungen benötigt man gelegentlich nur abgeschwächte Formen des Auswahlaxioms:

**Auswahlaxiom für abzählbare Mengen:** ($\mathrm{AC}_\omega$)
$a$ abzählbar $\wedge\, \forall x \in a\, x \neq \emptyset \to \exists f(D(f) = a \wedge \forall x \in a\, f(x) \in x)$,

**Axiom der abhängigen Auswahl:** (DC, *dependent choice*)
$Rel(R) \wedge a_0 \in a \;\wedge\; \forall x \in a \exists y \in a\, xRy$
$\to \exists f[f : \omega \longrightarrow a \;\wedge\; f(0) = a_0 \;\wedge\; \forall n < \omega\, f(n)Rf(n+1)]$

Es gilt:
AC $\to$ DC, DC $\to$ $\mathrm{AC}_\omega$ (aber die Umkehrungen sind nicht beweisbar).

Das $\mathrm{AC}_\omega$ wird vielfach in der Analysis benutzt (s.u.), insbesondere auch für die Aussage, daß die Vereinigung abzählbar-vieler abzählbarer Mengen wiederum abzählbar ist. DC benötigt man z. B. für die Äquivalenz verschiedener Endlichkeitsdefinitionen sowie zum Beweis, daß die Minimalitätsbedingung der Fundiertheit (s. (F1) von 6.1.1.3)

$$(F1) \quad \forall z(\emptyset \neq z \subseteq A \to \exists x \in z\, \forall y \in z\, y \not< x)$$

äquivalent ist zur Aussage, daß es für $<$ keine unendlich-absteigenden Ketten gibt:

$$(F11) \quad \neg \exists f(f : \omega \longrightarrow V \;\wedge\; \forall n < \omega\, f(n+1) < f(n))$$

(dabei benötigt man gerade das DC, um (F11) $\to$ (F1) zu zeigen, während die umgekehrte Richtung auch ohne DC beweisbar ist).

## 7.2 Anwendungen des Auswahlaxioms

### 7.2.1 Jeder Vektorraum besitzt eine Basis.

$K$ sei ein Körper, $V$ ein Vektorraum über dem Körper $K$, $V$ und $K$ seien Mengen (hier also $V$ nicht die Klasse aller Mengen!). Wir wenden zum Beweis das ZORNsche Lemma an und wählen

$$P := \{b \mid b \text{ Menge von linear unabhängigen Vektoren } \subseteq V\}, \leq = \subseteq \text{ auf} P.$$

Dann ist $\leq$ eine partielle Ordnung auf der Menge $P$, und für jede Kette $k \subseteq P$ ist $\bigcup k$ eine obere Schranke von $k$ in $P$. Nach dem ZORNschen Lemma existiert also ein maximales Element in $P$; eine maximale linear unabhängige Menge von Vektoren ist aber eine Basis.

**Bemerkungen**

1. A. Blass: *Existence of bases implies AC*, Contempory Math., Axiomatic Set Theory, AMS 31 (1983)) hat gezeigt, daß auch die Umkehrung gilt:
   Hat jeder Vektorraum eine Basis besitzt, so gilt das Auswahlaxiom.

2. Den Vektorraum $\mathbb{R}$ der reellen Zahlen kann man auch als (unendlich-dimensionalen) Vektorraum über den rationalen Zahlen $\mathbb{Q}$ auffassen, er besitzt dann also eine Basis, die man HAMEL-*Basis* nennt. Aus der Existenz einer solchen Basis folgt nun aber insbesondere die Existenz einer Funktion

   $$f : \mathbb{R} \longrightarrow \mathbb{R} \text{ mit } f(x+y) = f(x) + f(y) \text{ für alle reellen } x, y.$$

   Ein solches $f$ ist zwar $\mathbb{Q}$-linear, aber nicht $\mathbb{R}$-linear. Tatsächlich kann $f$ unstetig sein, und es ist sogar möglich, daß $W(f) \subseteq \mathbb{Q}$!

### 7.2.2 Der Satz von Hahn-Banach

(Formulierung und Beweis etwa bei Friedrichsdorf-Prestel, Kap. 10, pp.66ff)

### 7.2.3 Nicht-meßbare Mengen

*Es existiert eine Menge reeller Zahlen, die nicht Lebesgue-meßbar ist.* (VITALI 1905)

Die *Lebesgue-meßbaren* Mengen bilden eine Teilmenge $L \subseteq \mathcal{P}(\mathbb{R})$, auf welcher das *Lebesgue-Maß* definiert ist als Abbildung

$$m : L \longrightarrow \{x \mid x \in \mathbb{R} \wedge x \geq 0\} \cup \{\infty\},$$

so daß gilt:

(L1) $L$ ist ein *Mengenring*, d. h. $A, B \in L \rightarrow A \cup B, A \cap B, A - B \in L$,
und $L$ enthält die offenen und abgeschlossenen Intervalle:
$(a,b), [a,b] \in L$ für alle reellen Zahlen $a < b$.
Ferner ist $m([a,b]) = b - a$.

(L2) $m(A \cup B) = m(A) + m(B)$, falls $A \cap B = \emptyset$ *Additivität*

(L3) $A \subseteq B \rightarrow m(A) \leq m(B)$ *Monotonie*

(L4) Ist $(A_i | i < \omega)$ eine abzählbare Folge von Mengen $\in L$, so ist auch $\bigcup_{i<\omega} A_i \in L$, und falls $A_i \cap A_j = \emptyset$ für $i \neq j$, so ist

$m(\bigcup_{i<\omega} A_i) = \sum_{i<\omega} m(A_i)$. *$\sigma$-Additivität*

(L5) $A \in L \wedge r \in \mathbb{R} \rightarrow A + r = \{a + r \mid a \in A\} \in L$
und $m(A) = m(A + r)$ *Translationsinvarianz*

Auf den reellen Zahlen definieren wir eine Äquivalenzrelation durch

$$x \sim y :\leftrightarrow x, y \in [0,1] \ \wedge \ x - y \in \mathbb{Q}.$$

Nach dem Auswahlaxiom in der Form AC3´ existiert hierzu ein Repräsentantensystem $S$, d. h.

$$S \subseteq [0,1] \ \wedge \ \forall x \in [0,1] \ \exists! y \in S \ (x \sim y).$$

Setzen wir $S_r := \{x + r \mid x \in S\} = S + r$, so ist $\mathbb{R} = \bigcup_{s \in \mathbb{Q}} S_r$, und es ist

$$S_r \cap S_t = \emptyset \quad \text{für } r, t \in \mathbb{Q},\ r \neq t.$$

Wäre $S$ Lebesgue-meßbar, also $S \in L$, so erhielte man im Falle $m(S) = 0$: $m(\mathbb{R}) = 0$, und im Falle $m(S) > 0 : 2 = m([0,2]) = \infty$, Widerspruch! Somit ist die Menge $S$ nicht Lebesgue-meßbar. $\square$

### 7.2.4 Äquivalenz verschiedener Stetigkeitsdefinitionen

Auf den reellen Zahlen definiert man die übliche Topologie mit Hilfe der $\varepsilon$-*Umgebungen* von $x$:

$$U(\varepsilon,x) := (x-\varepsilon, x+\varepsilon) = \{y \in \mathbb{R} \mid x-\varepsilon < y < x+\varepsilon\},$$

sowie

(1) für $A \subseteq \mathbb{R}$ die *abgeschlossene Hülle* von $A$ durch

$$\begin{aligned} x \in \overline{A} : \quad &\leftrightarrow \quad \forall \varepsilon > 0 \quad U(\varepsilon,x) \cap A \neq \emptyset \quad \text{bzw.} \\ &\leftrightarrow \quad x = lim_{n<\infty} x_n \quad \text{für eine Folge} (x_n)_{n<\omega} \text{ mit } \forall n < \omega \; x_n \in A. \end{aligned}$$

(2) Es sei $f : \mathbb{R} \longrightarrow \mathbb{R}$, $x \in \mathbb{R}$. Dann definiert man
$f$ *stetig in* $x$:

$$\begin{aligned} &\leftrightarrow \quad \forall \varepsilon > 0 \, \exists \delta > 0 \, \forall y \in \mathbb{R} \; (|x-y| < \delta \rightarrow |f(x)-f(y)| < \varepsilon) \quad \text{bzw.} \\ &\leftrightarrow \quad f(x) = lim_{n<\infty} f(x_n) \text{ für alle Folgen } (x_n)_{n<\omega} \text{ mit } x = lim_{n<\infty} x_n. \end{aligned}$$

Um die Äquivalenz der jeweiligen Definitionen zu zeigen, benötigt man (in jeweils einer Richtung) das Auswahlaxiom - tatsächlich genügt hier das $\mathrm{AC}_\omega$.

### 7.2.5 Satz von Tychonoff

Das Produkt quasi-kompakter Räume ist quasi-kompakt. (Diese Aussage ist äquivalent zum Auswahlaxiom.)

### 7.2.6 Boolesches Primidealtheorem

Eine **Boolesche Algebra** ist ein Modell $(B, \wedge, \vee, -, 0, 1)$ der BOOLEschen Gesetze, die wir für die Aussagenlogik in 1.2.6 formuliert haben. Um den trivialen Fall auszuschließen, fügen wir noch $\emptyset \neq V$ als Axiom hinzu, so daß stets $0 \neq 1$ gilt und die kleinste BOOLEschen Algebra aus zwei Elementen besteht.

#### Beispiele und Bemerkungen

1. Für eine nicht-leere Menge M ist $(\mathcal{P}(M), \cap, \cup, -, \emptyset, M)$ eine BOOLEsche Algebra, die *Potenzmengenalgebra* von M. Eine Unteralgebra davon, definiert also auf einer Teilmenge $B \subseteq \mathcal{P}(M)$, die $\emptyset$ und $M$ enthält und abgeschlossen ist unter den Mengenoperationen $\cap, \cup, -$, heißt *Mengenalgebra*.

2. Jede endliche BOOLEschen Algebra ist isomorph zu einer Potenzmengenalgebra und hat damit $2^n$ Elemente für eine natürliche Zahl $n$. Die Menge $\{x \subseteq \mathbb{N} \mid x \textit{ endlich oder } \mathbb{N} - x \textit{ endlich}\}$ mit den üblichen mengentheoretischen Operationen ist eine abzählbare BOOLEsche Algebra, die aber nicht zu einer Potenzmengenalgebra isomorph sein kann.

3. Es gibt zahlreiche Beispiele BOOLEscher Algebren, die keine Mengenalgebren sind. Diese lassen sich am besten mit topologischen Hilfsmitteln beschreiben; aber es gibt auch natürliche Beispiele aus der Logik:

### Die Lindenbaum-Algebra einer Theorie $\mathsf{T}$

Für eine Theorie $\mathsf{T}$ der Sprache $\mathcal{L}$ definieren wir auf der Menge der Sätze von $\mathcal{L}$ eine Äqivalenzrelation durch

$$\varphi \sim_{\mathsf{T}} \psi :\iff \mathsf{T} \vdash \varphi \leftrightarrow \psi,$$
$$\mathcal{L}(\mathsf{T}) := \{[\sigma]_{\mathsf{T}} \mid \sigma\ \mathcal{L}\text{-Satz}\}$$

sei die Menge der zugehörigen Äquivalenzklassen. Den aussagenlogischen Operationen entsprechen dann Operationen auf $\mathcal{L}(\mathsf{T})$:

$$[\varphi]_{\mathsf{T}} \wedge [\psi]_{\mathsf{T}} := [\varphi \wedge \psi]_{\mathsf{T}}$$

und ähnlich für die übrigen Operationen. Damit wird $\mathcal{L}(\mathsf{T})$ zu einer BOOLEschen Algebra mit

$$0 = \{\varphi \mid \mathsf{T} \vdash \neg\varphi\} \text{ und } 1 = \{\varphi \mid \mathsf{T} \vdash \varphi\}.$$

Eine Theorie $\mathsf{T}$ ist somit vollständig und konsistent gdw $\mathcal{L}(\mathsf{T})$ isomorph zur 2-elementigen BOOLEschen Algebra ist.

Ein **Filter** einer BOOLEsche Algebra $(B, \wedge, \vee, -, 0, 1)$ ist eine nichtleere Teilmenge $F \subseteq B$ mit folgenden Eigenschaften:

(U1) $x, y \in F \Rightarrow x \wedge y \in F$,

(U2) $x \in F,\ y \in B \Rightarrow x \vee y \in F$.

Ein Filter $F$ heißt **echt** gdw $F \neq B$. Ein **Primfilter ( Ultrafilter)** ist ein echter Filter $F$ mit der zusätzlichen Eigenschaft

(U) $x \in F$ oder $B - x \in F$.

Ultrafilter kann man auch als maximale echte Filter charakterisieren; sie haben außerdem folgende Eigenschaften:

$$\begin{array}{rcl} x \notin U & \iff & B - x \in U \\ x \in U \text{ und } y \in U & \iff & x \wedge y \in U \\ x \in U \text{ oder } y \in U & \iff & x \vee y \in F \end{array}$$

und stehen damit in enger Verbindung mit dem Wahrheitsprädikat:

Das **Boolesche Primidealtheorem (BPI)** besagt:
*Jeder echte Filter in einer* BOOLE*schen Algebra läßt sich zu einem Ultrafilter erweitern,* (bzw.: *Jedes echte Ideal läßt sich zu einem Primideal erweitern.*)

(Die Begriffe **Ideal** und **Primideal** sind die zu **Filter** und **Ultrafilter** dualen Begriffe: vertausche $\vee$ mit $\wedge$.)

Das BPI folgt offensichtlich aus dem Auswahlaxiom (mittels des ZORNschen Lemmas), ist aber echt schwächer.

### Satz

*Das* BPI *ist äquivalent mit folgenden Aussagen:*

(i) *Jede* BOOLE*sche Algebra besitzt einen Ultrafilter.*

(ii) STONEscher Repräsentationssatz:
*Jede* BOOLE*sche Algebra ist isomorph zu einer Mengenalgebra.*

(iii) *Satz von* TYCHONOFF *für $T_2$-Räume*

(iv) *Vollständigkeitssatz für formale Sprachen erster Stufe*

(v) *Kompaktheitssatz für formale Sprachen erster Stufe*

Weitere Abschwächungen des Auswahlaxioms sind:

**Ordnungserweiterungsprinzip (OE):** *Jede partielle Ordnung läßt sich zu einer linearen Ordnung erweitern.*

**Ordnungsprinzip (OP):** *Jede Menge besitzt eine lineare Ordnung.*

Es gilt: AC $\Rightarrow$ BPI $\Rightarrow$ OE $\Rightarrow$OP (ohne Umkehrungen).

Weitere Ergebnisse und eine Übersicht findet man in dem Buch

G.H. MOORE: *Zermelo´s Axiom of Choice.* Springer 1982.

# Kapitel 8

# Mächtigkeiten und Kardinalzahlen

## 8.1 Endliche und abzählbare Mengen

Der Grundbegriff der Theorie der Kardinalzahlen ist die *Mächtigkeit* einer Menge. Zunächst definieren wir, wann zwei Mengen *gleichmächtig* sind:

$$a \sim b :\leftrightarrow \exists f(f : a \longleftrightarrow b) \qquad \textbf{gleichmächtig}$$

Es gibt verschiedene Möglichkeiten, den Begriff des *Endlichen* zu erfassen; da man hierbei auch das *Unendliche* definiert, ist der Begriff keineswegs trivial, was sich auch darin zeigt, dass die verschiedenen Definitionen oft erst mit Hilfe des Auswahlaxioms als äquivalent nachgewiesen werden können (s. (iv) des folgenden Satzes).

1. Der übliche Begriff der *endlichen* Menge greift auf die natürlichen Zahlen zurück:

Eine Menge $a$ ist endlich gdw sie mit Hilfe von natürlichen Zahlen bis zu einer Zahl $n$ abgezählt werden kann. Da nach unserer Festlegung $n = \{0, 1, \ldots, n-1\}$ ist, so haben wir eine besonders einfache Definition:

$$\begin{aligned} a \quad \text{endlich}: \quad &\leftrightarrow \quad \exists n < \omega \ (a \sim n), \\ a \quad \text{unendlich}: \quad &\leftrightarrow \quad \neg\, a \text{ endlich}. \end{aligned}$$

Ist $a$ endlich, so ist die natürliche Zahl $n$ mit $a \sim n$ (die man durch Abzählen bestimmt) eindeutig festgelegt und gibt die *Anzahl* der Elemente von $a$ an - für unendliche Mengen braucht dies aber nicht mehr zu gelten (siehe (i) des folgenden Satzes)!

2. Von WHITEHEAD-RUSSELL stammt die folgende Definition, die den Zahlbegriff vermeidet:
Eine Menge $u \subseteq \mathcal{P}(a)$ heißt *induktive Familie von Teilmengen von a* gdw

$$\emptyset \in u \wedge \forall x \in u\, \forall y \in a\; x \cup \{y\} \in u,$$

d.h. $u$ enthält die leere Menge und mit jeder Menge $x$ auch die um ein Element aus $a$ erweiterte Menge. Definiere:

$$a \text{ endlich} :\leftrightarrow \forall u \subseteq \mathcal{P}(a)(u \text{ induktive Familie von Teilmengen von } a \rightarrow a \in u).$$

Dieser Begriff ist äquivalent zu obiger Definition.

3. Die natürlichen Zahlen sind gleichmächtig mit einer echten Teilmenge (z. B. den geraden Zahlen). Diese paradoxe Eigenschaft wählt DEDEKIND gerade zur Charakterisierung unendlicher Mengen:

$$a \textit{ D-unendlich} :\leftrightarrow \exists x \subset a\, (a \sim x)$$

**Satz**

(i) $n \sim m \;\wedge\; n, m \in \omega \rightarrow n = m,$ *allgemeiner:*
$\alpha \sim m \;\wedge\; m \in \omega \rightarrow \alpha = m,$ *(jedoch:* $\omega \sim \omega + 1 \sim \omega + \omega$)

(ii) $\alpha$ *endlich* $\leftrightarrow \alpha \in \omega,$ *insbesondere*:
$\forall n \in \omega\, (n \textit{ endlich}) \;\wedge\; \omega \textit{ unendlich}$
$\forall n \in \omega\, \neg(n \textit{ D-unendlich}) \wedge \omega \textit{ D-unendlich}$

(iii) $a \textit{ D-unendlich} \leftrightarrow \exists x \subseteq a (x \sim \omega)$

(iv) $a \textit{ D-unendlich} \rightarrow a \textit{ unendlich},$
$a \textit{ D-unendlich} \leftrightarrow a \textit{ unendlich}$ (mit *DC* !)

*Beweis* von (iii) (DEDEKIND 1888): Sei $a$ D-unendlich, also
$f : a \longleftrightarrow b \subset a$ für ein $f$ und ein $b$. Wähle $a_0 \in a - b$ und setze $a_{n+1} = f(a_n)$ (rekursive Definition!).
Es ist dann $a_n \neq a_m$ für $n \neq m$ und somit $\{a_n \mid n < \omega\}$ eine abzählbar-unendliche Teilmenge von $a$.

Ist umgekehrt $\{a_n \mid n < \omega\}$ eine abzählbar-unendliche Teilmenge von $a$, wobei $a_n \neq a_m$ für $n \neq m$, so definieren wir eine Funktion $g : a \longrightarrow a$ durch

$$g(x) = \begin{cases} a_{n+1}, & \text{falls } x = a_n \\ x, & \text{sonst} \end{cases}$$

Es ist dann $g : a \longleftrightarrow a - \{a_0\} \subset a$. □

Eine D-unendliche Menge besitzt also eine abzählbar-unendliche Teilmenge; falls man das Auswahlaxiom (bzw. das Axiom DC) nicht voraussetzt, kann es unendliche Mengen geben, die keine abzählbar-unendliche Teilmenge enthalten!

**Eigenschaften endlicher Mengen**

(i) $\emptyset$ *und* $\{a\}$ *sind endlich.*

(ii) *Sind a und b endlich, so auch* $a \cup b,\ a \cap b,\ a - b,\ a \times b,\ \mathcal{P}(a),\ F[a],\ {}^a b$.

(iii) $a$ *endlich* $\wedge \forall x \in a\ (x$ *endlich*$) \rightarrow \bigcup a$ *endlich,*

(iv) $a$ *endlich* $\wedge\ b \subseteq a \rightarrow b$ *endlich.*

(v) **Induktionsprinzip für endliche Mengen:**

$$\varphi(\emptyset) \wedge \forall x[x \textit{ endlich} \wedge \varphi(x) \rightarrow \forall y(\varphi(x \cup \{y\}))] \rightarrow \forall x(x \textit{ endlich } \rightarrow \varphi(x)).$$

Diese Eigenschaften beweist man am einfachsten mit Hilfe des Endlichkeitsbegriffes von WHITEHEAD-RUSSELL, zeigt dann, dass unter diesem Begriff alle natürlichen Zahlen endlich sind und somit beide Endlichkeitsdefinitionen zusammenfallen. □

## 8.2 Abzählbare Mengen

Abzählbare Mengen lassen sich mit Hilfe der natürlichen Zahlen abzählen:

$a$ **abzählbar:** $\longleftrightarrow a = \emptyset \vee a = \{a_n \mid n < \omega\}$ *für eine Folge* $(a_n \mid n < \omega)$.

Damit sind dann auch die endlichen Mengen einbegriffen, da in der Abzählung Elemente mehrfach aufgezählt werden können. Wir unterscheiden:

1. $f : \omega \twoheadrightarrow a$ heißt **Aufzählung** von $a$ (es ist dann $a = \{f(n) \mid n < \omega\}$),

2. $f : \omega \leftrightarrow a$ heißt **Aufzählung** von $a$ **ohne Wiederholungen**.

Eine Aufzählung ohne Wiederholungen (also eine bijektive Abbildung auf die natürlichen Zahlen) liefert eine abzählbar-unendliche Menge:

$$a \textbf{ abzählbar-unendlich: } \longleftrightarrow a \sim \omega.$$

**Eigenschaften abzählbarer Mengen**

(i) *a abzählbar* $\leftrightarrow$ *a endlich* $\vee$ *a abzählbar-unendlich* $\leftrightarrow \exists \alpha \leq \omega\,(a \sim \alpha)$,

(ii) *Sind a und b abzählbar, so auch*

$$a \cup b,\ a \cap b,\ a - b,\ a \times b,\ F[a],$$
$$\mathcal{P}_{<\omega}(a) := \{x \mid x \subseteq a \wedge x\ endlich\},$$
$$a^{<\omega} := \{f \mid \exists n < \omega f : n \longrightarrow a\}.$$

(iii) *Unter der Voraussetzung des Auswahlaxioms* ($\mathsf{AC}_\omega$ *genügt*):
*a abzählbar* $\wedge\ \forall x \in a$ (*x abzählbar*) $\rightarrow \bigcup a$ *abzählbar*,

(iv) *a abzählbar* $\wedge\ b \subseteq a \rightarrow$ *b abzählbar.*

*Beweis* von (i): Sei zunächst $a$ abzählbar, also $f : \omega \twoheadrightarrow a$ für ein $f$, wobei wir $a$ als unendlich voraussetzen können. Zu zeigen ist: $a \sim \omega$, d. h. $a$ ist auch ohne Wiederholungen aufzählbar:
Setze $a_0 = f(0)$ und definiere (durch Rekursion) $a_{n+1} = f(k)$, wobei $k$ minimal ist mit $f(k) \neq a_i$ für alle $i < n+1$. (Da $a$ unendlich ist, muß ein solches $k$ existieren.) Es ist dann $n \mapsto a_n$ eine Bijektion von $\omega$ auf $a$.

Umgekehrt sind offenbar endliche wie auch abzählbar-unendliche Mengen abzählbar.

Wir zeigen wir nun (iii), wobei wir annehmen können, daß die betrachteten Mengen nicht-leer sind: Dazu geben wir zunächst eine **Paarfunktion** an:

$$f : \omega \times \omega \longleftrightarrow \omega, \qquad f(n,m) = 2^n \cdot (2m+1) - 1.$$

Sei nun $a = \{a_n \mid n < \omega\}$ abzählbar und $\forall n \in \omega$ ($a_n$ abzählbar), also
$\forall n \in \omega\ \exists g (g : \omega \twoheadrightarrow a_n)$.
Nach $\mathsf{AC}_\omega$ existiert eine Folge $(g_n)_{n\in\omega}$ mit $\forall n \in \omega\ a_n = \{g_n(m) \mid \in \omega\}$, und somit

$$\bigcup a = \bigcup_{n\in\omega} a_n = \{g_n(m) \mid m,n \in \omega\} = \{g(n,m) \mid m,n \in \omega\},$$

wobei $g : \omega \times \omega \twoheadrightarrow \bigcup_{n\in\omega} a_n$ definiert ist durch $g(n,m) = g_n(m)$.

Die gewünschte Abbildung $h : \omega \twoheadrightarrow \bigcup_{n\in\omega} a_n$ erhält man nun aus $g$ durch Vorschalten der Inversen $f^{-1}$ der oben definierten Paarfunktion $f$.

(ii) beweist man ähnlich, wobei man sich leicht überzeugt, daß man in diesem Fall das abzählbare Auswahlaxiom nicht benötigt. □

## 8.3 Überabzählbare Mengen

Während die Potenzmenge einer endlichen Menge auch wieder endlich ist, ist die Potenzmenge einer abzählbar-unendlichenen Menge nicht mehr abzählbar, also **überabzählbar**: Wäre

$$\begin{array}{rl} \mathcal{P}(\omega) = \{a_n \mid n \in \omega\}, & \text{so wäre auch} \\ a := \{m \in \omega \mid m \notin a_m\} = a_n & \text{für ein } n, \text{dann aber} \\ n \in a_n \leftrightarrow n \notin a_n & \text{Widerspruch!} \end{array}$$

Wir erhalten also einen Widerspruch wie im Falle der RUSSELLschen Antinomie! In ähnlicher Weise kann man zeigen, daß die Menge aller zahlentheoretischen Funktionen und auch die Menge aller reellen Zahlen $\mathbb{R}$ überabzählbar sind.

## 8.4 Mächtigkeiten

Obwohl wir die *Mächtigkeit* einer Menge noch nicht erklärt haben, konnten wir Mengen als gleichmächtig definieren, wenn sie sich eineindeutig aufeinander abbilden lassen. Ebenso lassen sich Mengen hinsichtlich ihrer Größe vergleichen, ohne sie vorher aufzählen zu müssen:

$$\begin{array}{rlll} a \sim b: & \leftrightarrow & \exists f (f : a \longleftrightarrow b) & a \text{ ist } \textbf{gleichmächtig} \text{ mit } b \\ a \preccurlyeq b: & \leftrightarrow & \exists f (f : a \rightarrowtail b) & a \text{ ist } \textbf{kleiner oder gleichmächtig} \text{ mit } b \\ & \leftrightarrow & \exists x \subseteq b\ (a \sim x) & (a \text{ ist } \textit{schmächtiger} \text{ als } b) \\ a \prec b: & \leftrightarrow & a \preccurlyeq b \wedge a \not\sim b & a \text{ ist } \textbf{kleiner} \text{ als } b \end{array}$$

(Genauer sollte man für $a \preccurlyeq b$ sagen: *a ist von Mächtigkeit kleiner oder gleich b*.) So ist z.B.

$$n \preccurlyeq \omega, n \prec \omega, \omega \preccurlyeq \omega+1, \omega+1 \preccurlyeq \omega, \omega \prec \mathcal{P}(\omega), \text{ aber } \omega+1 \not\prec \omega, \omega \not\prec \omega+1.$$

**Lemma**

| | | |
|---|---|---|
| $(i)$ | $a \preccurlyeq a$ | reflexiv |
| $(ii)$ | $a \preccurlyeq b \wedge b \preccurlyeq c \rightarrow a \preccurlyeq c$ | transitiv |
| $(iii)$ | $a \subseteq b \rightarrow a \preccurlyeq b$ | aber i. a. nicht umgekehrt! |
| $(iv)$ | $a \sim c \wedge a \preccurlyeq b \rightarrow c \preccurlyeq b$ | ebenso für $\prec$ |
| $(v)$ | $b \sim d \wedge a \preccurlyeq b \rightarrow a \preccurlyeq d$ | ebenso für $\prec$ |

Somit ist die Gleichmächtigkeit (Äquivalenz) von Mengen eine Äquivalenzrelation und eine Kongruenzrelation bezüglich $\preccurlyeq$ und $\prec$; die Relation $\preccurlyeq$ ist reflexiv und transitiv. Der folgende Satz besagt, daß $\preccurlyeq$ antisymmetrisch (bis auf $\sim$) ist, was trivial ist, wenn man das Auswahlaxiom benutzt. Als eines der wenigen Ergebnisse der Theorie der Mächtigkeiten läßt er sich aber auch ohne diese Voraussetzung beweisen:

### 8.4.1 Satz von Cantor-Schröder-Bernstein

$$a \preccurlyeq b \ \wedge \ b \preccurlyeq a \ \rightarrow a \sim b$$

*Beweis:* Wir führen die Behauptung zunächst auf den einfacheren Fall

$$(*) \quad a \preccurlyeq b \ \wedge \ b \subseteq a \ \rightarrow a \sim b$$

zurück: Nach Voraussetzung existieren Abbildungen

$$\begin{array}{ll} h : a \longleftrightarrow a' \subseteq b, \ g : b \longleftrightarrow b' \subseteq a, & \text{also} \\ f := g \circ h : a \longleftrightarrow a_1 \subseteq b' & \text{mit} \quad a_1 := g[a'] \sim a' \sim a, \text{ d. h.} \\ a_1 \subseteq b' \subseteq a \ \wedge \ a_1 \sim a \ \wedge \ b \sim b', & \text{somit} \\ a \preccurlyeq b' \ \wedge \ b' \subseteq a \ \wedge \ b' \sim b. & \end{array}$$

Nach (*) gilt dann $a \sim b'$, also auch $a \sim b$, wie zu zeigen war.

Zum Beweis von (*) gehen wir von einer injektiven Abbildung

$$f : a \rightarrowtail b \subseteq a$$

aus und konstruieren daraus eine Abbildung

$$g : a \leftrightarrow b$$

wie folgt:

$$g(x) = \begin{cases} f(x), & \text{falls } x \in \bigcup_{n\in\omega} f^n[a-b] \\ x & \text{sonst.} \end{cases}$$

Dabei ist $f^0(y) = y, f^{n+1}(y) = f(f^n(y))$ (numerische Rekursion). Es gilt nun:

(i) $g$ ist surjektiv, d. h. $W(g) = b$:

Sei $d \in b$. Falls $d \in \bigcup_{n\in\omega} f^n[a-b]$, so ist $d \in f^n[a-b]$ für ein $n$, und zwar $n > 0$ wegen $d \in b$. Dann ist aber $d = f(f^{n-1}(y))$ für ein $y$ und damit $d \in W(g)$. Im anderen Fall ist aber $d = g(d)$ und damit auch $d \in W(g)$.

(ii) $g$ ist injektiv:

Da $f$ injektiv ist, so auch $g$ auf $\bigcup_{n\in\omega} f^n[a-b]$, und als identische Abbildung ist sie auch auf dem Komplement injektiv. Ist aber einerseits $x \in \bigcup_{n\in\omega} f^n[a-b]$ und andererseits $y \in a - \bigcup_{n\in\omega} f^n[a-b]$, so muß auch $g(x) \neq g(y)$ sein, da $g(x) \in \bigcup_{n\in\omega} f^n[a-b]$, während $g(y) = y$ im Komplement liegt. Somit ist $g$ eine Bijektion von $a$ auf $b$. □

Der folgende Satz benötigt zum Beweis notwendig das Auswahlaxiom, da er hierzu äquivalent ist:

### 8.4.2 Vergleichbarkeitssatz von Hartogs

$$a \preccurlyeq b \vee b \preccurlyeq a$$

*Beweis:* Nach dem Wohlordnungssatz gibt es Ordinalzahlen $\alpha, \beta$ mit $a \sim \alpha$ und $b \sim \beta$. Da Ordinalzahlen vergleichbar sind (und zwar $\alpha \subseteq \beta \vee \beta \subseteq \alpha$), überträgt sich diese Beziehung in der Form $a \preccurlyeq b \vee b \preccurlyeq a$ auf die entsprechenden Mengen. □

### Folgerung

*Jede unendliche Menge enthält eine abzählbar-unendliche Teilmenge (und ist damit D-unendlich).*

Daß es zu jeder Menge eine mit größerer Mächtigkeit gibt, zeigt der

### 8.4.3 Satz von Cantor

$$a \prec \mathcal{P}(a)$$

*Beweis:* Da durch $x \mapsto \{x\}$ eine injektive Funktion definiert wird, ist $a \preccurlyeq \mathcal{P}(a)$. Die Annahme $a \sim \mathcal{P}(a)$ widerlegt man wie im Fall $a = \omega$ durch ein Diagonalargument: Falls $f : a \twoheadrightarrow \mathcal{P}(a)$, so setze man $d := \{x \in a \mid x \notin f(x)\}$. Dann erhält man wegen der Surjektivität von $f$ ein $b \in a$ mit $d = f(b)$, was aber mit
$b \in f(b) = d \leftrightarrow b \notin b$ zum Widerspruch führt! □

**Alternative zum Größenvergleich**

Da bei einer Abbildung Mengen möglicherweise auf dasselbe Element abgebildet werden können, wird man das Bild einer Menge $a$ als höchstens so groß wie $a$ selbst ansehen. Man findet daher auch folgende Definition:

$$a \preccurlyeq^* b :\leftrightarrow a = \emptyset \vee \exists f (f : b \twoheadrightarrow a).$$

Für die Definition der Abzählbarkeit haben wir z. B.:

$$a \text{ abzählbar} \leftrightarrow a \preccurlyeq^* \omega \leftrightarrow a \preccurlyeq \omega,$$

und der obige Satz von CANTOR besagt gerade: $a \not\preccurlyeq^* \mathcal{P}(a)$. Da eine injektive Funktion eine Umkehrfunktion besitzt, so gilt stets

$$a \preccurlyeq b \to a \preccurlyeq^* b,$$

aber für die Umkehrung benötigt man das Auswahlaxiom (oder zumindest eine Wohlordnung von $b$)! Ebenso gilt zwar stets $F[a] \preccurlyeq^* a$, aber $F[a] \preccurlyeq a$ allgemein nur, wenn das Auswahlaxiom vorausgesetzt ist.

## 8.5 Kardinalzahlen

Für die *Mächtigkeit* einer Menge $a$, bezeichnet mit $\overline{\overline{a}}$, soll gelten:

$$a \sim b \leftrightarrow \overline{\overline{a}} = \overline{\overline{b}}.$$

Diese Eigenschaft kann man natürlich leicht erfüllen, indem man (nach FREGE-RUSSELL) $\overline{\overline{a}}$ als zugehörige Äquivalenzklasse $\{x \mid x \sim a\}$ definiert, allerdings ist diese dann stets eine echte Klasse (außer für $a = \emptyset$). Die Anzahl einer endlichen Menge kann man aber auch dadurch bestimmen, daß man sie abzählt: Ist

$$a = \{a_0, \ldots, a_{n-1}\} \text{ mit } a_i \neq a_j \text{ für } i \neq j,$$

so ist $n$ die Anzahl der Elemente von $a$.

Ist jedoch $a$ eine unendliche Menge, so können folgende Probleme auftauchen:

1. $a$ besitzt keine Aufzählung,

2. $a$ besitzt Aufzählungen verschiedener Länge.

Das erste Problem läßt sich vermeiden, indem man das Auswahlaxiom voraussetzt: Dann läßt sich jede Menge wohlordnen:

$$a = \{a_\xi \mid \xi < \alpha\} \text{ für ein } \alpha,$$

aber das zweite Problem bleibt bestehen: Selbst bei injektiven Aufzählungen ist die ”Länge” $\alpha$ nicht eindeutig bestimmt. Als *Kardinalzahl* von $a$ wählt man daher das kleinstmögliche derartige $\alpha$:

$$|a| := \overline{\overline{a}} := \mu\alpha(a \sim \alpha) \qquad \textbf{Kardinalzahl} \text{ von } a$$

**Bemerkung**

Auch ohne Voraussetzung des Auswahlaxioms kann man (etwa mit Hilfe der VON-NEUMANNschen Hierarchie) Mächtigkeiten definieren, die Mengen sind. In diesem Fall nennt man *Kardinalzahlen* die Mächtigkeiten von wohlgeordneten Mengen. Allerdings läßt sich (wie wir schon bei obigen Sätzen gesehen haben) ohne Auswahlaxiom keine befriedigende Theorie der Mächtigkeiten aufbauen. Daher werden wir im folgenden das Auswahlaxiom stets voraussetzen und nicht zwischen Kardinalzahlen und Mächtigkeiten unterscheiden.

**Beispiele**

$$|n| = n \quad \text{für jede natürliche Zahl } n,$$
$$|\mathbb{N}| = |\omega| = |\omega + 1| = |\omega + 2| = \ldots = |\omega + \omega| = \omega.$$

Tatsächlich ist jede unendliche Kardinalzahl eine Limeszahl (aber nicht umgekehrt), die Kardinalzahlen bilden somit eine echte Teilklasse *Cn* der Ordinalzahlen. *Cn* ist trotzdem immer noch sehr groß: Das Supremum einer Menge von Kardinalzahlen ist wieder eine Kardinalzahl, und nach dem Satz von CANTOR gibt es keine größte Kardinalzahl.

### 8.5.1 Operationen auf den Kardinalzahlen

Die arithmetischen Operationen auf den Kardinalzahlen definiert man analog zum endlichen Fall, wobei im Falle der Addition zu beachten ist, daß man als Summe zweier Kardinalzahlen die Vereinigung *disjunkter* Mengen der entsprechenden Kardinalzahl wählt:

$$\begin{aligned} \kappa \oplus \lambda &= |(\kappa \times \{0\}) \cup (\lambda \times \{1\})| \\ \kappa \odot \lambda &= |\kappa \times \lambda| \\ \kappa^{\underline{\lambda}} &= |{}^{\lambda}\kappa| = |\{f \mid f : \lambda \to \kappa\}| \end{aligned}$$

Endliche Ordinal- und Kardinalzahlen stimmen überein (es sind gerade die natürlichen Zahlen), und für diese Fälle stimmen diese Operationen mit den entsprechenden Operationen auf den Ordinalzahlen ebenfalls überein. Für unendliche Kardinalzahlen ergeben sich aber wesentliche Unterschiede. So sind die Operationen der Addition und Multiplikation auf den Kardinalzahlen (wie im Falle der natürlichen Zahlen, aber im Gegensatz zu den ordinalen Operationen) kommutativ, assoziativ und distributiv - wenngleich diese Gesetze trivial sind; es gilt nämlich der

### 8.5.2 Satz von Hessenberg

*Es seien $\kappa, \lambda$ unendliche Kardinalzahlen. Dann gilt:*

$$\kappa \oplus \lambda = \kappa \odot \lambda = \max\{\kappa, \lambda\}.$$

(Tatsächlich braucht nur eine der beiden Zahlen unendlich zu sein, im Falle der Multiplikation darf aber natürlich keine = 0 sein.) Dieses Ergebnis läßt sich (mit einfachen Monotoniegesetzen) zurückführen auf

### Satz

*Für unendliche Kardinalzahlen gilt:*

$$\kappa \odot \kappa = \kappa.$$

Damit vereinfacht sich die Bestimmung der Kardinalzahl unendlicher Mengen in vielen Fällen, insbesondere für die Mengen:

$$\begin{aligned} a^{<\omega} : &= \{f \mid \exists n < \omega\, f : n \to a\} \\ \mathcal{P}_{<\omega}(a) : &= \{x \subseteq a \mid x\, endlich\} \end{aligned}$$

### Satz

*Für* $\emptyset \neq b \preccurlyeq a$, *a unendlich, gilt:*

(i) $|a \cup b| = |a \times b| = |a|$,

(ii) $|a^{<\omega}| = |\mathcal{P}_{<\omega}(a)| = |a|$.

Zur Aufzählung der unendlichen Kardinalzahlen benutzt man seit CANTOR den hebräischen Buchstaben $\aleph$ (aleph):

$$\omega = \aleph_0 < \aleph_1 < \ldots \aleph_\omega < \ldots$$

$\aleph_0$ ist also die Kardinalzahl abzählbar-unendlicher Mengen, $\aleph_1$ die erste überabzählbare Kardinalzahl. Während Addition und Multiplikation unendlicher Kardinalzahlen trivial sind, ist schon der Wert der einfachsten Potenz unbekannt:

Die Mächtigkeit der Menge $\mathbb{R}$ der reellen Zahlen ist $2^{\aleph_0}$ (kardinale Potenz!), und da $\mathbb{R}$ überabzählbar ist, so gilt $\aleph_1 \leq 2^{\aleph_0}$. Die

**Cantorsche Kontinuumshypthese** CH $\quad \aleph_1 = 2^{\aleph_0}$

ist aus den Axiomen von $\mathsf{ZF}$ + AC weder beweisbar (COHEN 1963) noch widerlegbar (GÖDEL 1938). Trotzdem gibt es einige einfache Abschätzungen:

### 8.5.3 Satz von Bernstein

$$\begin{aligned} 2 \leq \kappa \leq \aleph_\alpha \quad &\rightarrow \quad 2^{\aleph_\alpha} = \kappa^{\aleph_\alpha} = \aleph_\alpha^{\aleph_\alpha} \\ \beta \leq \alpha \quad &\rightarrow \quad \aleph_\alpha < 2^{\aleph_\alpha} = \aleph_\beta^{\aleph_\alpha} = \aleph_\alpha^{\aleph_\alpha} \end{aligned}$$

Wir definieren eingeschränkte Potenzmengenoperationen:

$$\begin{aligned} \mathcal{P}_{<\lambda}(a) :&= \{x \mid x \subseteq a \wedge |x| < \lambda\}, \\ \mathcal{P}_{\leq\lambda}(a) :&= \{x \mid x \subseteq a \wedge |x| \leq \lambda\}, \\ \mathcal{P}_{\lambda}(a) :&= \{x \mid x \subseteq a \wedge |x| = \lambda\}. \end{aligned}$$

### Korollar

*Für unendliches* $a, \kappa, \lambda$ :

(i) $|\mathcal{P}_{<\omega}(a)| = |a|$,

(ii) $|\{f \mid Fkt(f) \wedge D(f) \subseteq \kappa \wedge D(f) \textit{ endlich} \wedge W(f) \subseteq \lambda\}| = \max(\kappa, \lambda)$.

(iii) $|\mathcal{P}_\lambda(\kappa)| = |\mathcal{P}_{\leq\lambda}(\kappa)| = \kappa^\lambda$ für $\lambda \leq \kappa$, insbesondere:

(iv) $|\mathcal{P}_\kappa(\kappa)| = |\mathcal{P}(\kappa)| = \kappa^\kappa = 2^\kappa$.

# Teil III

# Grundlagen der Modelltheorie

# Kapitel 9

# Hin und her: Vollständigkeit und elementare Äquivalenz

Mathematische Theorien lassen sich grob in zwei Klassen unterteilen:

1. Theorien, die Eigenschaften *eines bestimmten* Modells beschreiben wollen (z. B. die Theorie der natürlichen (reellen, komplexen) Zahlen oder die Theorie der Euklidischen Ebene) sowie andererseits

2. Theorien, welche gemeinsame Eigenschaften *vieler* Modelle zusammenfassen, wobei sie aber durchaus zulassen, daß Modelle sich bezüglich anderer Eigenschaften unterscheiden (wie die meisten algebraischen Theorien: Gruppen, Ringe, Körper, Vektorräume, ..., aber auch die Theorie topologischer Räume, Graphen und Ordnungen).

## 9.1 Elementare Äquivalenz

Die Theorie $\mathsf{Th}(\mathcal{A})$ einer Struktur $\mathcal{A}$ unterscheidet sich von allgemeinen Theorien wie folgt:

- $\mathsf{Th}(\mathcal{A}) \models \sigma \iff \mathcal{A} \models \sigma$,
- $\mathsf{T} \models \sigma \iff$ *für <u>alle</u>* $\mathcal{A} \models \mathsf{T} : \mathcal{A} \models \sigma$,

die Gültigkeit eines Satzes $\sigma$ in $\mathsf{Th}(\mathcal{A})$ hängt also nur von der Gültigkeit in dem *einen* Modell $\mathcal{A}$ ab, sonst von der Gültigkeit in *allen* Modellen der Theorie. Somit ist $\mathsf{Th}(\mathcal{A})$ *semantisch vollständig* in folgendem Sinne:

$$\mathsf{T}\ \textbf{(semantisch) vollständig} :\iff \text{für alle } \mathcal{L}\text{-Sätze } \sigma : \mathsf{T} \models \sigma \text{ oder } \mathsf{T} \models \neg\sigma.$$

**Achtung:** Für *Formeln* $\varphi$ gilt i. a. **nicht**: $\mathcal{A} \models \varphi$ oder $\mathcal{A} \models \neg\varphi$, denn sonst müßte der universelle Abschluß von $\varphi$ oder der universelle Abschluß von $\neg\varphi$ in $\mathcal{A}$ gelten (aber in einer nicht-kommutativen Gruppe $\mathcal{G}$ gilt weder $\forall x \forall y\ (x+y=y+x)$ noch $\forall x \forall y\ (x+y \neq y+x)$). Dagegen gilt nach der Wahrheitsdefinition für *Sätze* $\sigma$ stets

$$\mathcal{A} \models \sigma \quad \text{oder} \quad \mathcal{A} \models \neg\sigma.$$

Dagegen gilt i. a. auch **nicht**: $\mathsf{T} \models \sigma$ oder $\mathsf{T} \models \neg\sigma$, denn der Satz $\sigma$ könnte in einem Modell von $\mathsf{T}$ gelten, in einem anderen aber nicht (wie das Kommutativgesetz im Falle der Gruppentheorie). Dagegen gilt die letzte Alternative, wenn sich je zwei Modelle von $\mathsf{T}$ in ihren elementaren Eigenschaften (ausgedrückt durch Sätze der zugehörigen Sprache) nicht unterscheiden:

Sind $\mathcal{A}$ und $\mathcal{B}$ Strukturen derselben Sprache $\mathcal{L}$, so heißen sie **elementar-äquivalent** genau dann, wenn sie dieselben Sätze erfüllen:

$$\boxed{\mathcal{A} \equiv \mathcal{B} :\iff \text{für alle } \mathcal{L}\text{-Sätze } \sigma :\ \mathcal{A} \models \sigma \iff \mathcal{B} \models \sigma.}$$

**Beispiele und Bemerkungen**

1. $(\mathbb{N},<) \equiv (\mathbb{N}\backslash\{0\},<)$, aber $(\mathbb{N},<,1) \not\equiv (\mathbb{N}\backslash\{0\},<,1)$.

2. $(\mathbb{Q},+,\cdot,0,1) \not\equiv (\mathbb{R},+,\cdot,0,1)$ als Körper (2 ist in $\mathbb{Q}$ kein Quadrat), betrachten wir sie jedoch als geordnete Strukturen, so gilt:
   $(\mathbb{Q},<) \equiv (\mathbb{R},<)$ (wird später bewiesen), dagegen $(\mathbb{Z},<) \not\equiv (\mathbb{Q},<)$.

3. $\mathcal{A} \equiv \mathcal{B} \iff \mathsf{Th}(\mathcal{A}) = \mathsf{Th}(\mathcal{B})$.

4. **Eine Theorie ist vollständig gdw je zwei ihrer Modelle elementar-äquivalent sind.**

5. Hat $\mathsf{T}$ überhaupt ein Modell, so ist $\mathsf{T}$ vollständig gdw $\mathsf{T} = \mathsf{Th}(\mathcal{A})$ für ein (oder: alle) Modelle $\mathcal{A}$ von $\mathsf{T}$.

6. Für einen Isomorphismus $h : \mathcal{A} \cong \mathcal{B}$ gilt:

$$\mathcal{A} \models \varphi[a_1, \ldots a_n] \iff \mathcal{B} \models \varphi[h(a_1), \ldots h(a_n)],$$

   d. h. ein Isomorphismus ist eine **elementare** Einbettung (s.u.), und somit sind isomorphe Strukturen äquivalent. Das obige Beispiel in (2) zeigt, daß die Umkehrung aber nicht zu gelten braucht: elementar-äquivalente Strukturen können sogar verschiedene Mächtigkeiten haben und lassen sich daher nicht einmal als Mengen bijektiv aufeinander abbilden.

7. Ist $\mathcal{A}$ endlich, so ist im Falle $\mathcal{A} \equiv \mathcal{B}$ auch $\mathcal{B}$ endlich, und zwar von derselben Anzahl (da beide dieselben Anzahlformeln erfüllen). Tatsächlich kann man zeigen, daß $\mathcal{A}$ und $\mathcal{B}$ in diesem Fall sogar isomorph sind!

Die obigen Beispiele zeigen zudem, daß elementare Äquivalenz in starkem Maße von der Sprache abhängt, die man zugrunde legt. Die Relation der elementaren Äquivalenz ist ein wichtiger Grundbegriff der Modelltheorie, läßt sich aber i. a. schwer nachweisen, da man über Sätze einer Sprache keine Induktion führen kann. (Meistens weist man diese Relation nach über die stärkere Isomorphiebeziehung oder zeigt, daß beide Strukturen Modelle einer vollständigen Theorie sind.)

## 9.2 Die Theorie der dichten linearen Ordnung

Wir wollen hier ein Musterbeispiel einer vollständigen Theorie behandeln. Dazu werden wir zeigen, daß je zwei abzählbare Modelle isomorph sind; aus dem später folgenden Kriterium 10.6 von VAUGHT wird sich dann ergeben, daß diese Theorie vollständig ist.

Die Theorie LO der *linearen Ordnung* wird in einer Sprache mit einem 2-stelligen Relationszeichen $<$ formuliert und hat folgende Axiome:

| | | |
|---|---|---|
| (L1) | $\forall x\, x \not< x$ | **irreflexiv** |
| (L2) | $\forall x \forall y \forall z\, (x < y \wedge y < z \rightarrow x < z)$ | **transitiv** |
| (L3) | $\forall x \forall y\, (x < y \vee x = y \vee y < x)$ | **connex** |

Für dichte Ordnungen gilt zusätzlich

| | | |
|---|---|---|
| (D) | $\forall x \forall y \exists z\, (x < y \rightarrow x < z \wedge z < y)$ | **dicht** |

Diese Theorie ist noch nicht vollständig; sie wird es aber, wenn man Aussagen über die Endpunkte hinzunimmt, etwa

(OE) $\forall x \exists y\,(x < y) \wedge \forall x \exists y\,(y < x)$ **ohne Endpunkte**

DLO sei die Theorie der dichten linearen Ordnungen (Axiome: L1−L3, D); Beispiele von Modellen von DLO + OE sind $(\mathbb{Q},<)$ und $(\mathbb{R},<)$, ersteres ist bis auf Isomorphie das einzige abzählbare Modell:

**Isomorphiesatz von Cantor**

*Je zwei abzählbare dichte lineare Ordnungen ohne Endpunkte sind isomorph.*

*Beweis:* Es seien $(A,<),(B,<')$ abzählbare dichte Ordnungen ohne Endpunkte, $A = \{a_0,a_1,\ldots\}, B = \{b_0,b_1,\ldots\}$. Wir konstruieren einen Isomorphismus $h : (A,<) \cong (B,<')$ mit $a'_n \mapsto b'_n$ wie folgt:

hin: $a_0 \mapsto b_0$, wir setzen also $a'_0 := a_0, b'_0 := b_0$.

her: In der Aufzählung $B = \{b_0,b_1,\ldots\}$ kommt nach $b_0$ als nächstes Element $b_1 =: b'_1$ an die Reihe: wir suchen in der Aufzählung $A = \{a_0,a_1,\ldots\}$ ein entsprechendes $a'_1$, welches in der Ordnung $<$ zu $a'_0 = a_0$ so liegt, wie $b'_1$ zu $b'_0$ (da die Ordnungen keine Endpunkte haben, ist dies möglich).

⋮

hin: Es sei $n = 2i-1$ ungerade und die Zuordnung $a'_0 \mapsto b'_0,\ldots,a'_n \mapsto b'_n$ bereits gefunden. Wir wählen $a'_{n+1}$ als das erste $a$ in der Aufzählung $A = \{a_0,a_1,\ldots\}$, welches wir noch nicht nach $B$ abgebildet haben (welches also unter den $a'_0,\ldots a'_n$ noch nicht vorgekommen ist) und ordnen diesem ein entsprechendes $b'_{n+1}$ in $B$ zu, welches unter den bisherigen Bildern $b'_0,\ldots,b'_n$ nicht vorkommt, so daß es zu diesen in der Ordnung $<'$ so liegt, wie $a$ zu den $a'_0,\ldots a'_n$, mithin

$$(\{a'_0,\ldots,a'_{n+1}\},<) \cong (\{b'_0,\ldots,b'_{n+1}\},<),$$

was möglich ist, da die Ordnung $<'$ auf $B$ dicht und ohne Endpunkte ist.

her: Es sei $n = 2i$ gerade und die die Zuordnungen $a'_0 \mapsto b'_0,\ldots,a'_n \mapsto b'_n$ bereits gefunden. Wir wählen jetzt umgekehrt ein $b'_{n+1}$ als das erste $b$ in der Aufzählung $B = \{b_0,b_1,\ldots\}$, für welches wir noch kein Urbild in $A$ gefunden haben (welches also unter den $b'_0,\ldots b'_n$ noch nicht vorgekommen ist) und ordnen diesem ein entsprechendes $a'_{n+1}$ in $A$ zu, welches unter den bisherigen Urbildern $a'_0, \ldots, a'_n$ nicht vorkommt, so daß es zu diesen in der Ordnung $<$ so liegt, wie $b$ zu den $b'_0, \ldots, b'_n$, mithin

$$(\{a'_0, \ldots, a'_{n+1}\}, <) \cong (\{b'_0, \ldots, b'_{n+1}\}, <),$$

was möglich ist, da die Ordnung $<$ auf $A$ dicht und ohne Endpunkte ist.

Damit erhalten wir eine ordnungstreue bijektive Abbildung $h : A \longleftrightarrow B$, die wegen der "hin"-Fälle schließlich alle Elemente von $A$ abbildet und wegen der "her"-Fälle auch alle Elemente von $B$ als Bilder annimmt. □

**Bemerkungen**

1. Auf ähnliche Weise kann man zeigen, daß auch zwei abzählbare dichte lineare Ordnungen mit erstem und letztem Element isomorph sind (indem man diese zuerst aufeinander abbildet und dann mit dem Hin- und Herverfahren fortfährt), ähnlich bei anderen Festlegungen über die Endelemente. Die (zunächst unvollständige) Theorie der dichten linearen Ordnungen hat also 4 verschiedene Vervollständigungen: ohne Endelemente, mit beiden Endpunkten, mit nur kleinstem und ohne größtes sowie mit größtem, aber ohne kleinstes Element.

2. Die Methode des "Hin und Her" stammt von HUNTINGTON 1904; CANTOR hat nur den "Hin"-Teil benutzt, welcher übrigens zeigt, daß man $(\mathbb{Q}, <)$ in jede dichte lineare Ordnung einbetten kann. (Diese Methode läßt sich in der Modelltheorie sehr gut verallgemeinern.)

3. Überabzählbare dichte lineare Ordnungen derselben Mächtigkeit brauchen jedoch nicht isomorph zu sein: Ist etwa $(\mathbb{R}, <)$ eine solche Ordnung, so ist die Ordnung, die man erhält, wenn man eine Kopie der abzählbaren Struktur $(\mathbb{Q}, <)$ anfügt, von derselben Mächtigkeit, aber nicht isomorph. Es gilt sogar: Nimmt man aus den rationalen Zahlen einen Punkt heraus, so ist die verbleibende Struktur immer noch eine abzählbare lineare Ordnung ohne Endpunkte (ähnlich, wenn man endlich-viele Punkte herausnimmt), anders jedoch im Falle von $(\mathbb{R}, <)$:

$$\begin{aligned} & (\mathbb{Q}\setminus\{0\}, <) \cong (\mathbb{Q}, <), \\ \text{aber} \quad & (\mathbb{R}\setminus\{0\}, <) \ncong (\mathbb{R}, <). \end{aligned}$$

Die Unterschiede beruhen darauf, daß $(\mathbb{Q}, <)$ sehr viele "Lücken" hat, die erst durch Hinzunahme der irrationalen Zahlen geschlossen werden; dagegen ist $(\mathbb{R}, <)$ so dicht (ein *Kontinuum*), daß dort jede monoton aufsteigende beschränkte Folge einen Limes hat.

## 9.3 Kategorizität

Definiert man eine Theorie $\mathsf{T}$ als **kategorisch** genau dann, wenn je zwei Modelle von $\mathsf{T}$ isomorph sind, so sind zwar vollständige Theorien mit einem endlichen Modell kategorisch, aber - wie wir im nächsten Kapitel sehen werden - keine Theorie mit einem unendlichen Modell. Man kann dann höchstens noch danach fragen, ob Modelle derselben Mächtigkeit isomorph sind:

$\mathsf{T}$ **$\kappa$-kategorisch** $:\iff$ *je zwei Modelle von $\mathsf{T}$ der Kardinalität $\kappa$ sind isomorph.*

Nach dem Isomorphiesatz von CANTOR ist die Theorie $\mathsf{DLO}$ der dichten linearen Ordnung ohne Endpunkte $\aleph_0$-kategorisch, woraus leicht folgt, daß sie auch vollständig ist. Diese Folgerung werden wir im nächsten Kapitel verallgemeinern: der Test von VAUGHT 10.6 liefert mittels der Kategorizität ein Kriterium für die Vollständigkeit einer Theorie.

Es gibt Theorien, die

- für keine unendliche Kardinalzahl $\kappa$ $\kappa$-kategorisch sind (z. B. die Theorie der reell-abgeschlossenen Körper),
- für $\kappa = \aleph_0$, aber für keine überabzählbare Kardinalzahl $\kappa$ $\kappa$-kategorisch sind (z. B. die Theorie der dichten linearen Ordnung ohne Endpunkte),
- nur für überabzählbare Kardinalzahl $\kappa$ $\kappa$-kategorisch sind (z. B. die Theorie der algebraisch-abgeschlossenen Körper),
- für alle unendlichen Kardinalzahlen $\kappa$ $\kappa$-kategorisch sind (z.B. die Theorie der unendlichen Mengen in der Sprache mit $=$ und den Konstanten $c_n$; einzige Axiome sind die Sätze $c_n \neq c_m$ für $n \neq m$).

Mehr Unterscheidungen gibt es nicht: Nach einem Satz von MORLEY (1965) gilt:

*Ist eine Theorie* $\mathsf{T}$ *(in einer abzählbaren Sprache)* $\kappa$*-kategorisch für* <u>*ein*</u> *überabzählbares* $\kappa$*, so ist* $\mathsf{T}$ $\kappa$*-kategorisch für* <u>*alle*</u> *überabzählbaren* $\kappa$*.*

(Entsprechend verallgemeinert gilt nach SHELAH dieses Ergebnis auch für Theorien in einer überabzählbaren Sprache.) Interessant ist außerdem das Ergebnis von VAUGHT, daß eine vollständige Theorie niemals genau *zwei* abzählbare nicht-isomorphe Modelle besitzen kann. Eine Reihe neuerer Arbeiten (insbesondere von SHELAH) beschäftigt sich mit der Frage

- wieviele nicht-isomorphe Modelle der Mächtigkeit $\kappa$ kann eine Theorie $\mathsf{T}$ besitzen und

- wie kann man diese Modelle klassifizieren?

# Kapitel 10

# Auf und ab: Elementare Substrukturen

## 10.1 Elementare Substrukturen

$\mathcal{A}$ und $\mathcal{B}$ seien Strukturen derselben Sprache $\mathcal{L}$, $\mathcal{A} \subseteq \mathcal{B}$. Dann heißt $\mathcal{A}$ **elementare Substruktur** von $\mathcal{B}$ gdw beide bezüglich der Elemente der kleineren Struktur dieselben Eigenschaften erfüllen:

$$\mathcal{A} \preccurlyeq \mathcal{B} :\iff \text{ für alle } \mathcal{L}\text{-Formeln } \varphi(v_1, \dots v_n) \text{ und für alle } a_1, \dots a_n \in A :$$

$$\begin{aligned} \mathcal{A} \models \varphi[a_1, \dots a_n] \quad &\iff \quad \mathcal{B} \models \varphi[a_1, \dots a_n], \quad \text{d. h.} \\ \mathcal{A} \preccurlyeq \mathcal{B} \quad &\iff \quad \mathcal{A}_A \equiv \mathcal{B}_A. \end{aligned}$$

Ein Homomorphismus $h : \mathcal{A} \longrightarrow \mathcal{B}$ heißt eine **elementare Einbettung** gdw für alle $\mathcal{L}$-Formeln $\varphi(v_1, \dots v_n)$ und für alle $a_1, \dots a_n \in A$ gilt:

$$\mathcal{A} \models \varphi[a_1, \dots a_n] \iff \mathcal{B} \models \varphi[h(a_1), \dots h(a_n)].$$

Indem man für $\varphi$ die Formeln $v_1 = v_2$ bzw. $R(v_1, \dots v_n)$ wählt, sieht man, daß in diesem Fall $h$ tatsächlich eine Einbettung sein muß. Die obige Bedingung besagt also gerade, daß das Bild $h[\mathcal{A}]$ eine elementare Substruktur von $\mathcal{B}$ ist; durch Identifizieren kann man dann also $\mathcal{A} \preccurlyeq \mathcal{B}$ erreichen.

Diese Begriffe kann man auch mit Hilfe des Diagramms charakterisieren; der Beweis des früheren Diagrammlemmas 3.4.1 läßt sich nämlich leicht erweitern:

## 10.2 Diagrammlemma (Fortsetzung)

*Es sei $\mathcal{A}$ eine $\mathcal{L}$-Struktur, $(\mathcal{B},(\underline{a}^{\mathcal{B}})_{a\in A})$ eine $\mathcal{L}_A$-Struktur und $h : A \longrightarrow B$ definiert durch $h(a) = \underline{a}^{\mathcal{B}}$. Dann gilt:*

(i) $(\mathcal{B},(h(a))_{a\in A}) \models \mathsf{D}(\mathcal{A}) \iff h : \mathcal{A} \longrightarrow \mathcal{B}$ *ist eine Einbettung,*

(ii) $(\mathcal{B},(h(a))_{a\in A}) \models \mathsf{D}_{\mathcal{L}}(\mathcal{A}) \iff h : \mathcal{A} \longrightarrow \mathcal{B}$ *ist eine elementare Einbettung.*

*Insbesondere gilt für $A \subseteq B$:*

(iii) $\mathcal{B}_A \models \mathsf{D}(\mathcal{A}) \iff \mathcal{A} \subseteq \mathcal{B}$,

(iv) $\mathcal{B}_A \models \mathsf{D}_{\mathcal{L}}(\mathcal{A}) \iff \mathcal{A} \preccurlyeq \mathcal{B}$,

(v) $\mathcal{A}_A \equiv \mathcal{B}_A \iff \mathcal{A} \preccurlyeq \mathcal{B}$.

## 10.3 Kriterium von Tarski

*Es seien $\mathcal{A},\mathcal{B}$ $\mathcal{L}$-Strukturen, $\mathcal{A} \subseteq \mathcal{B}$. Dann sind äquivalent:*

(i) $\mathcal{A} \preccurlyeq \mathcal{B}$

(ii) *für alle $\mathcal{L}$-Formeln $\varphi(v_0,\dots,v_n)$ und beliebige $a_1,\dots a_n \in A$ gilt:*
$\mathcal{B} \models \exists v_0\varphi[a_1,\dots a_n] \Longrightarrow$ *es gibt ein $a_0 \in A$ mit $\mathcal{B} \models \varphi[a_0,a_1,\dots a_n]$.*

Der *Beweis* ist sehr einfach: $(i) \Rightarrow (ii)$ folgt unmittelbar aus der Definition einer elementaren Substruktur, während man die Umkehrung durch Induktion über den Formelaufbau beweist. Dabei sind alle Fälle trivial bis auf den Fall einer Formel der Gestalt $\exists v_o\varphi$, und hier liefert das Kriterium den entscheidenden Schritt von der Existenz in der Struktur $\mathcal{B}$ zur Existenz in der Unterstruktur $\mathcal{A}$. $\square$

Dieses Kriterium ist besonders nützlich, weil in (ii) nur die Gültigkeit in der Struktur $\mathcal{B}$ vorkommt. Für Anwendungen kann man auch oft den folgenden Satz benutzen:

### Satz

*Es sei $\mathcal{A} \subseteq \mathcal{B}$ und für beliebige $a_1,\dots a_n \in A, b \in B$ existiere ein Automorphismus*

$$j : \mathcal{B} \longrightarrow \mathcal{B} \text{ mit } j(a_i) = a_i \text{ und } j(b) \in A,$$

*welcher also die $a_i$ festläßt und $b$ nach $A$ abbildet. Dann ist $\mathcal{A} \preccurlyeq \mathcal{B}$.* $\square$

**Beispiele**

1. $(\mathbb{Q},<) \preccurlyeq (\mathbb{R},<)$ (wende obigen Satz an!) und somit auch $(\mathbb{Q},<) \equiv (\mathbb{R},<)$.
2. $\mathcal{A} \preccurlyeq \mathcal{B} \Longrightarrow \mathcal{A} = \mathcal{B}$, falls $A$ oder $B$ endlich sind. (Denn nach Voraussetzung gelten in $\mathcal{A}$ dieselben Anzahlformeln wie in $\mathcal{B}$.)

Endliche Strukturen haben somit weder echte elementare Substrukturen noch echte elementare Erweiterungen, anders dagegen im Falle unendlicher Strukturen: Diese haben in jeder kleineren Mächtigkeit $\kappa$ (solange noch $\kappa \geq card(\mathcal{L})$) eine elementare Substruktur:

## 10.4 Satz von Löwenheim-Skolem-Tarski (abwärts)

*$\mathcal{A}$ sei eine $\mathcal{L}$-Struktur, $\kappa$ eine Kardinalzahl mit card($\mathcal{L}$) $\leq \kappa \leq$ card(A). Dann existiert eine Struktur $\mathcal{B}$ mit*

$$\mathcal{B} \preccurlyeq \mathcal{A} \text{ und } card(B) = \kappa.$$

Zusatz: *Ist $A_0 \subseteq A$ beliebig mit $card(A_0) \leq \kappa$, so kann man zusätzlich $A_0 \subseteq B$ verlangen.*

*Beweis:* Es sei $A_0 \subseteq A$ gegeben mit $card(A_0) \leq \kappa \leq card(A)$. Indem wir $A_0$ nötigenfalls vergrößern, können wir annehmen, daß $card(A_0) = \kappa$ ist. Falls für Elemente $a_1, \ldots, a_n \in A_0$

$$(*) \quad \mathcal{A} \models \exists v_0 \varphi[a_1, \ldots, a_n]$$

gilt, so müssen wir ein (von $\varphi$ und den $a_1, \ldots, a_n$ abhängiges) $b_{\exists v_0 \varphi}(a_1, \ldots, a_n)$ zu $A_0$ hinzunehmen; $A_1$ sei die Menge, die aus $A_0$ entsteht, indem solche Beispiele (und zwar für alle möglichen Existenzformeln und alle möglichen Belegungen mit Elementen aus $A_0$) hinzugenommen werden. Nach unseren früheren Ergebnissen über unendliche Kardinalzahlenbleibt $card(A_1) = card(A_0) = \kappa$.

Nun kann aber (*) möglicherweise für eine Formel $\varphi$ und neue Elemente $a_1, \ldots, a_n \in A_1$ gelten, die nicht in $A_0$ enthalten sind; wir müssen also das Verfahren wiederholen:

Ist $A_i$ definiert, so ergänze man es um geeignete Elemente von $A$, so daß gilt:

$$\textit{Falls für } a_1, \ldots, a_n \in A_i : \mathcal{A} \models \exists v_0 \varphi[a_1, \ldots, a_n],$$
$$\textit{so existiert ein } a \in A_{i+1} \textit{ mit } \mathcal{A} \models \varphi[a, a_1, \ldots, a_n].$$

Wie im ersten Schritt läßt sich $A_{i+1}$ aus $A_i$ gewinnen, ohne daß die Kardinalität sich ändert. Schließlich setzen wir

$$B := \bigcup_{i \in \mathbb{N}} A_i$$

und erhalten eine Menge $B$ mit $A_0 \subseteq B \subseteq A$ und $card(B) = \kappa$. Beim ersten Schritt müssen wir die Konstanten von $\mathcal{A}$ bereits aufgenommen haben (mit (*) für die Formel $\exists v_0(v_0 = c)$) und in allen weiteren Schritten unter den Funktionen von $\mathcal{A}$ abgeschlossen haben (entsprechend der Formel $\exists v_0(v_0 = F(v_1, \ldots, v_n))$, so daß $B$ das Universum einer Unterstruktur $\mathcal{B}$ von $\mathcal{A}$ ist. Schließlich gilt $\mathcal{B} \preccurlyeq \mathcal{A}$, weil unsere Konstruktion gerade so eingerichtet ist, daß das Kriterium von TARSKI-VAUGHT anwendbar ist. □

Einfacher mit Hilfe des Kompaktheitssatzes erhalten wir den

## 10.5 Satz von Löwenheim-Skolem-Tarski (aufwärts)

*$\mathcal{A}$ sei eine unendliche $\mathcal{L}$-Struktur, $\kappa$ eine Kardinalzahl mit card(A), card($\mathcal{L}$) $\leq \kappa$. Dann existiert eine Struktur $\mathcal{B}$ mit*

$$\mathcal{A} \preccurlyeq \mathcal{B}\ ,\ card(B) = \kappa.$$

*Beweis:* Wir wählen zunächst eine Menge $C$ mit $A \subseteq C$ und $card(C) = \kappa$ und erweitern die Theorie von $\mathcal{A}$ (mit Hilfe neuer Konstanten) um Sätze, so daß jedes Modell mindestens so viele Elemente wie $C$ hat:

$$\mathsf{T}' = \mathsf{Th}(\mathcal{A}) \cup \{\underline{c} \neq \underline{d} | c, d \in C, c \neq d\}.$$

Nach dem Kompaktheitssatz hat $\mathsf{T}'$ ein Modell, etwa $\mathcal{B}$, in welchem die neuen Konstanten $\underline{c}$ durch Elemente von $B$ interpretiert werden, und zwar verschiedene Konstanten durch verschiedene Elemente von $B$. Indem wir zu einer isomorphen Struktur übergehen, können wir annehmen, daß $\underline{c}^{\mathcal{B}} = c$ ist und somit $C \subseteq B$, also $card(B) \geq \kappa$. Die Sprache von $\mathsf{T}'$ hat wegen $\kappa \geq card(\mathcal{L})$ die Mächtigkeit $\kappa$, also kann man auch annehmen, daß $card(B) = \kappa$ ist (indem man den Abwärts-Satz benutzt).

Schließlich gilt wegen $\mathcal{B} \models \mathsf{Th}(\mathcal{A}) : \mathcal{A} \equiv \mathcal{B}$. Die stärkere Aussage $\mathcal{A} \preccurlyeq \mathcal{B}$ erhält man, indem man dasselbe Argument benutzt, jedoch statt $\mathsf{Th}(\mathcal{A})$ das elementare Diagramm $\mathsf{Th}(\mathcal{A}_A)$ wählt, das aus allen in $\mathcal{A}$ wahren $\mathcal{L}_A$-Sätzen besteht. □

**Folgerungen**

1. *Ist $\mathcal{A}$ eine $\mathcal{L}$-Struktur und $\kappa \geq card(\mathcal{L})$ eine Kardinalzahl, so gibt es eine Struktur $\mathcal{B}$ mit $card(B) = \kappa$ und*

   - $\mathcal{B} \preccurlyeq \mathcal{A}$ im Falle $\kappa \leq card(A)$,
   - $\mathcal{A} \preccurlyeq \mathcal{B}$ im Falle $card(A) \leq \kappa$.

2. Insbesondere hat jede Theorie $\mathsf{T}$, die ein unendliches Modell besitzt, für *jede* Kardinalzahl $\kappa \geq card(\mathcal{L})$ ein Modell der Kardinalzahl $\kappa$.

3. Noch spezieller: Eine Theorie $\mathsf{T}$ in einer abzählbaren Sprache $\mathcal{L}$, die überhaupt ein Modell hat, hat ein abzählbares Modell. Dieser Satz von LÖWENHEIM (1915) ist eines der frühesten Ergebnisse der mathematischen Logik.

**Zusammenfassung**

Mit früheren Ergebnissen erhalten wir:

1. Hat eine vollständige Theorie ein endliches Modell, so sind also alle Modelle dieser Theorie isomorph (zu diesem einen Modell).

2. Hat eine Theorie beliebig große endliche Modelle, so nach 3.3.2 auch ein unendliches Modell.

3. Für eine Theorie $\mathsf{T}$ mit einem unendlichen Modell gibt es Modelle in jeder Mächtigkeit $\kappa \geq card(\mathcal{L})$, wobei $\mathcal{L}$ die Sprache der Theorie $\mathsf{T}$ ist.

## 10.6 Test von Vaught

*Ist eine Theorie $\mathsf{T}$ $\kappa$-kategorisch für ein $\kappa \geq card(\mathcal{L})$ und besitzt $\mathsf{T}$ nur unendliche Modelle, so ist $\mathsf{T}$ vollständig.*

*Beweis:* Gegeben seien zwei Modelle $\mathcal{A}, \mathcal{B}$ von $\mathsf{T}$, von denen wir zeigen müssen, daß sie elementar-äquivalent sind. Da sie unendlich sind, kann man nach den Sätzen von LÖWENHEIM-SKOLEM-TARSKI zu elementaren Erweiterungen (bzw. Substrukturen) $\mathcal{A}', \mathcal{B}'$ dieser Strukturen übergehen, die dann dieselbe Mächtigkeit $\kappa$ haben, also nach Voraussetzung isomorph, insbesondere elementar-äquivalent sind. Damit sind dann aber auch $\mathcal{A}, \mathcal{B}$ elementar-äquivalent. $\square$

# Kapitel 11

# Vereinigungen, Durchschnitte und Ketten

Bekanntlich ist der Durchschnitt von (affinen) Unterräumen eines Vektorraumes wieder ein Unterraum (oder die leere Menge), während die Vereinigungsmenge lediglich einen Unterraum aufspannt. Ähnlich ist es im allgemeinen Fall von Substrukturen einer Struktur $\mathcal{A}$:

### Durchschnitt

Es sei $(\mathcal{A}_i)_{i\in I}$ eine Familie von Substrukturen einer Struktur $\mathcal{A}$. Dann ist der Durchschnitt $\bigcap_{i\in I}|\mathcal{A}_i|$ entweder $\emptyset$ oder Grundbereich einer Substruktur von $\mathcal{A}$, die mit $\bigcap_{i\in I}\mathcal{A}_i$ bezeichnet wird.

Denn sind die Bedingungen (i) und (ii) von Satz 2.2.2 für alle $|\mathcal{A}_i|$ erfüllt, so auch für deren Durchschnitt. - Wichtiger als der Durchschnitt ist die

### Vereinigung

Es sei wieder $(\mathcal{A}_i)_{i\in I}$ eine Familie von Substrukturen einer Struktur $\mathcal{A}$. Dann *erzeugt* die Vereinigungsmenge $\bigcup_{i\in I}|\mathcal{A}_i|$ eine Substruktur von $\mathcal{A}$, die mit $\bigcup_{i\in I}\mathcal{A}_i$ bezeichnet wird, ihr Grundbereich ist $\overline{\bigcup_{i\in I}|\mathcal{A}_i|}$.

Interessanter ist der Fall, daß für eine vorgegebene Familie von Strukturen eine gemeinsame Oberstruktur nicht vorhanden ist, sondern erst gebildet werden soll. Dazu müssen aber die Strukturen bestimmte Verträglichkeitsbedingungen erfüllen.

## Ketten von Strukturen

Eine Familie $(\mathcal{A}_i)_{i\in I}$ von $\mathcal{L}$-Strukturen bildet eine *Kette*, wenn je zwei Strukturen miteinander vergleichbar sind in dem Sinne, daß von zwei Strukturen jeweils eine Struktur eine Substruktur der anderen ist:

$$(K) \quad (\mathcal{A}_i)_{i\in I} \textit{ ist eine } \textbf{Kette} : \iff \forall i,j \in I: \ \mathcal{A}_i \subseteq \mathcal{A}_j \textit{ oder } \mathcal{A}_j \subseteq \mathcal{A}_i.$$

Damit sind auch je endlich-viele Strukturen der Kette vergleichbar; insbesondere: wenn $\mathcal{A}_{i_1} \ldots \mathcal{A}_{i_n}$ Strukturen der Kette sind, so sind alle $\subseteq \mathcal{A}_{i_k}$ für ein $k$.

Später benötigen wir einen stärkeren Begriff:

$$(\mathcal{A}_i)_{i\in I} \textit{ ist eine } \textbf{elementare Kette} : \iff \forall i,j \in I: \mathcal{A}_i \preccurlyeq \mathcal{A}_j \textit{ oder} \mathcal{A}_j \preccurlyeq \mathcal{A}_i.$$

Die **Vereinigung** $\bigcup_{i\in I}\mathcal{A}_i$ einer Kette können wir wie folgt definieren:

(V1) $|\mathcal{A}| := \bigcup_{i\in I}|\mathcal{A}_i|$,

(V2) $R^{\mathcal{A}} := \bigcup_{i\in I} R^{\mathcal{A}_i}$ für jedes Relationszeichen $R$ der Sprache $\mathcal{L}$,

(V3) $f^{\mathcal{A}} := \bigcup_{i\in I} f^{\mathcal{A}_i}$ für jedes Funktionszeichen $f$ der Sprache $\mathcal{L}$,

(V4) $c^{\mathcal{A}} := c^{\mathcal{A}_i}$ für jede Individuenkonstante $c$ der Sprache $\mathcal{L}$.

Das bedeutet also für ein Relationszeichen $R$ und Elemente $a_1,\ldots,a_n \in A$:

$$R^{\mathcal{A}}(a_1,\ldots,a_n) \iff R^{\mathcal{A}_i}(a_1,\ldots,a_n) \text{ für ein } i \in I \text{ mit } a_1,\ldots,a_n \in A_i,$$

wobei es wegen (K) auf die Auswahl des gewählten $i$ nicht ankommt, ähnlich im Falle der Funktionszeichen. Die Individuenkonstanten werden wegen (K) in allen $\mathcal{A}_i$ durch dasselbe Element interpretiert (und damit auch in der Vereinigung).

## 11.1 Satz über Ketten von Strukturen

*Ist $\mathcal{A} = \bigcup_{i\in I}\mathcal{A}_i$ die Vereinigung der Kette $(\mathcal{A}_i)_{i\in I}$, so gilt:*

(i) *$\mathcal{A}$ ist eine $\mathcal{L}$-Struktur mit $\mathcal{A}_i \subseteq \mathcal{A}$ für jedes $i \in I$,*

(ii) *$\mathcal{A}$ ist die kleinste derartige $\mathcal{L}$-Struktur.*

(iii) *Ist* $(\mathcal{A}_i)_{i\in I}$ *eine elementare Kette, so gilt* $\mathcal{A}_i \preccurlyeq \mathcal{A}$ für jedes $i \in I$. (**Kettenlemma von Tarski**)

In den meisten Fällen (wie auch in den folgenden Beispielen) kann man sich auf **aufsteigende $\omega$-Ketten** von Strukturen beschränken; in diesen Fällen ist die Indexmenge $I = \omega = \mathbb{N}$ und es gilt:

$$\forall n < m \quad \mathcal{A}_n \subseteq \mathcal{A}_m.$$

**Beispiele**

1. Es sei $A_n := [1/2^n, 1]$ das abgeschlossene Intervall der reellen Zahlen, $\mathcal{A}_n = (A_n, \leq)$ mit der gewöhnlichen $\leq$-Beziehung auf den reellen Zahlen. Dann ist die Vereinigung dieser Strukturen das halb-offene Intervall $(A, \leq)$ mit $A = (0, 1]$.

2. Es sei $A_n := \{x \in \mathbb{Z} | -n \leq x \leq n\}$, $\mathcal{A}_n = (A_n, \leq)$ mit der gewöhnlichen $\leq$-Beziehung auf den ganzen Zahlen. Dann ist die Vereinigung dieser Strukturen $(\mathbb{Z}, \leq)$.

In beiden Fällen ist die Vereinigung von (dichten) linearen Ordnungen wieder eine (dichte) lineare Ordnung, aber die Existenz von Endpunkten bleibt nicht erhalten! Im ersten Fall haben wir isomorphe, insbesondere elementar-äquivalente Strukturen, aber die Vereinigung hat neue Eigenschaften. Man kann nun aber genau angeben, welche Eigenschaften erhalten bleiben:

$$\sigma \text{ ist ein } \forall\exists\text{-Satz:} \iff \sigma = \forall x_1 \ldots \forall x_n \exists y_1 \ldots \exists y_m\, \varphi$$
$$\text{für eine quantorenfreie Formel } \varphi$$

**Beispiele**

1. $\forall x \exists y\, (x \leq y \wedge x \neq y)$ (Unbeschränktheit nach oben) ist ein $\forall\exists$-Satz, nicht aber $\exists x \forall y\, x \leq y$ (Existenz eines kleinsten) bzw. $\exists x \forall y\, y \leq x$ (Existenz eines größten Elementes).

2. Wichtige Sätze der Algebra sind ebenfalls $\forall\exists$-Sätze: $\forall x \exists y\, (x \circ y = e)$ (Existenz eines inversen Elementes) sowie
$$\forall x_0 \ldots x_n \exists y\, (y^{n+1} + x_n y^n + \ldots + x_0 = 0)$$
(Existenz von Nullstellen für normierte Polynome vom Grad $n+1$).

$\forall\exists$-Sätze bleiben bei der Vereinigung von Ketten erhalten:

**Erhaltungssatz für Ketten**

*Ist $\mathcal{A} = \bigcup_{i\in I} \mathcal{A}_i$ die Vereinigung der Kette $(\mathcal{A}_i)_{i\in I}$, so gilt für jeden $\forall\exists$-Satz $\sigma$:*

$$\mathcal{A}_i \models \sigma \text{ für alle } i \in I \implies \mathcal{A} \models \sigma.$$

*Beweis:* Es sei $\sigma = \forall x_1 \ldots \forall x_n \exists y_1 \ldots \exists y_m\ \varphi$ für eine quantorenfreie Formel $\varphi$ und $\sigma$ gelte in allen $\mathcal{A}_i$. Um die Gültigkeit in $\mathcal{A}$ nachzuweisen, seien beliebige $a_1, \ldots,\ a_n \in A$ gegeben. Nach Definition der Vereinigungsmenge einer Kette können wir ein $i \in I$ finden, so daß alle $a_1, \ldots, a_n \in A_i$ sind. Weil $\sigma$ in $\mathcal{A}_i$ gilt, gibt es $b_1, \ldots, b_m \in A_i$ mit

$$\mathcal{A}_i \models \varphi[a_1, \ldots, a_n, b_1, \ldots, b_m].$$

Da $\varphi$ quantorenfrei ist, sich also aussagenlogisch aus atomaren Formeln zusammensetzt, und da $\mathcal{A}_i \subseteq \mathcal{A}$, so gilt die Formel $\varphi$ mit derselben Belegung auch in $\mathcal{A}$. Die $b_1, \ldots, b_m$ sind aber auch in $A$, so daß dort der Existenzsatz gilt:

$$\mathcal{A} \models \exists y_1 \ldots \exists y_m\ \varphi[a_1, \ldots, a_n, b_1, \ldots, b_m]. \qquad \square$$

Das obige Ergebnis ist optimal: Eine Theorie heiße **induktiv** gdw sie abgeschlossen ist unter Vereinigungen von Ketten, d. h. ist $(\mathcal{A}_i)_{i\in I}$ eine Kette von Modellen von $\mathsf{T}$, so ist auch die Vereinigung $\bigcup_{i\in I} \mathcal{A}_i$ ein Modell von $\mathsf{T}$.

## 11.2 Satz von Chang- Łoś-Szusko

*Eine Theorie $\mathsf{T}$ ist induktiv gdw sie ein Axiomensystem aus $\forall\exists$-Sätzen besitzt.*

Beim Beweis (den wir hier nicht ausführen können) wird für die Induktivität von $\mathsf{T}$ nur die Abgeschlossenheit unter aufsteigenden $\omega$-Ketten benötigt, hieraus folgt dann also schon die Abgeschlossenheit unter beliebigen Ketten!

Zahlreiche algebraische Theorien sind induktiv: die Theorie der (ABELschen) Gruppen, der Ringe, der (reell- bzw. algebraisch-abgeschlossenen) Körper, der Vektorräume, etc.

# Kapitel 12

# Produkte und Ultraprodukte

Während sich die Vereinigung einer Familie von Strukturen nur unter geeigneten Verträglichkeitsbedingungen sinnvoll definieren läßt, kann man das *(direkte) Produkt* stets bilden. Dazu erinnern wir an den Begriff der *Familie* von Mengen:

*$(a_i)_{i\in I}$ ist die auf der Indexmenge $I$ definierte*
*Funktion $f$ mit $f(i) = a_i$ für alle $i \in I$.*

Das mengentheoretische *Produkt* einer Familie $(A_i)_{i\in I}$ ist eine Verallgemeinerung des endlichen Cartesischen Produktes:

$$\prod_{i\in I} A_i := \{(a_i)_{i\in I} | \, \forall i \in I \; a_i \in A_i\} \quad \textbf{Produkt} \text{ der Mengen } (A_i)_{i\in I}$$

Das *Auswahlaxiom* (s. 7.2) besagt gerade, daß die Produktmenge nicht-leerer Mengen wiederum nicht-leer ist.

**Bezeichnungen:** Um endliche Folgen von Elementen aus der Produktmenge zu bezeichnen, werden wir Querstriche und obere Indizes verwenden:

- $a^1 = (a_i^1)_{i\in I}, \dots, a^n = (a_i^n)_{i\in I}$
  seien $n$ Elemente der Produktmenge $A = \prod_{i\in I} A_i$,
- $\overline{a} = (a^1, \dots, a^n)$ bezeichnet das entsprechende $n$-Tupel aus $A^n$,
- $\overline{a}_i = (a_i^1, \dots, a_i^n)$ die $i$-te *Projektion* des $n$-Tupels $\overline{a} = (a^1, \dots, a^n)$.

## 12.1 Direktes Produkt

$(\mathcal{A}_i)_{i\in I}$ sei eine Familie von $\mathcal{L}$-Strukturen, $I \neq \emptyset$. Das direkte Produkt der $(\mathcal{A}_i)_{i\in I}$ ist die $\mathcal{L}$-Struktur $\mathcal{A} = \prod_{i\,\in I} \mathcal{A}_i$, die wie folgt erklärt ist:

(P1) $A := \prod_{i\,\in I} A_i$ ist der Grundbereich von $\mathcal{A}$,

(P2) $R^{\mathcal{A}}(\overline{a}) :\Longleftrightarrow \forall i \in I\ R^{\mathcal{A}_i}(\overline{a}_i)$ für jedes Relationszeichen $R$ von $\mathcal{L}$,

(P3) $f^{\mathcal{A}}(\overline{a}) := (f^{\mathcal{A}_i}(\overline{a}_i))_{i\in I}$ für jedes Funktionszeichen $f$ von $\mathcal{L}$,

(P4) $c^{\mathcal{A}} := (c^{\mathcal{A}_i})_{i\in I}$ für jede Individuenkonstante $c$ von $\mathcal{L}$.

Für eine endliche Indexmenge $I = \{1, \ldots, m\}$ schreibt man auch $\mathcal{A}_1 \times \ldots \times \mathcal{A}_m$, und sind alle Strukturen dieselbe Struktur $\mathcal{A}$, so bezeichnet man das direkte Produkt auch mit $\mathcal{A}^I$ und nennt es **direkte Potenz** von $\mathcal{A}$.

**Beispiele**

1. Ist $\mathcal{R}$ der gewöhnliche *Vektorraum* über den reellen Zahlen, so ist $\mathcal{R}^n$ das $n$-fache direkte Produkt (die $n$-te Potenz) von $\mathcal{R}$, denn die Operationen werden nach P3 komponentenweise erklärt. Allgemeiner ist das direkte Produkt von Gruppen wieder eine Gruppe.

2. Das direkte Produkt $\mathcal{A} \times \mathcal{A}$ der *linearen Ordnungen* $\mathcal{A} = (\mathbb{R}, <)$ ist jedoch nur eine partielle Ordnung, denn nach P2 werden die Relationen im Produkt ebenfalls komponentenweise erklärt, was in diesem Fall aber bedeutet, daß etwa die Paare $(0,1)$ und $(1,0)$ in der Ordnungsrelation des Produktes unvergleichbar sind. Ebenso ist das direkte Produkt des *Körpers* der reellen Zahlen mit sich kein Körper, da im Produkt Nullteiler auftauchen: $(0,1) \cdot (1,0) = (0,0)$ bei komponentenweiser Multiplikation!

Da sich nicht alle Eigenschaften von *allen* Strukturen auf das direkte Produkt übertragen lassen, wird man versuchen, die Eigenschaften von *sehr vielen* Strukturen zu erhalten. Der Begriff von "sehr vielen" Elementen von $I$ (oder von einer "sehr großen" Teilmenge von $I$) wird durch den Begriff des *Filters* umschrieben:

## 12.2 Filter und Ultrafilter

Es sei $I \neq \emptyset$ eine Menge und $\mathcal{P}(I) := \{X \mid X \subseteq I\}$ ihre **Potenzmenge**. Eine Menge $F \subseteq \mathcal{P}(I)$ heißt (echter) **Filter auf** $I$ gdw

(F1) $\emptyset \notin F$ und $I \in F$,

(F2) $A \in F \textit{ und } A \subseteq B \subseteq I \Longrightarrow B \in F$,

(F3) $A, B \in F \Longrightarrow A \cap B \in F$.

Für einen **Ultrafilter auf** $I$ gilt zusätzlich für alle $A \subseteq I$:

(F4) $(entweder)\, A \in F$ oder $I \backslash A = \{x \in I \mid x \notin A\} \in F$.

Ultrafilter haben besonders schöne Eigenschaften, die sie für die Logik interessant machen: *Ist U Ultrafilter auf I, so gilt für alle* $A, B \subseteq I$:

(U1) $A \notin U \iff I \backslash A \in U$,

(U2) $A \in U \wedge B \in U \iff A \cap B \in U$,

(U2) $A \in U \vee B \in U \iff A \cup B \in U$.

**Beispiele und Bemerkungen**

1. Für jedes $\emptyset \neq A \subseteq I$ ist $F(A) := \{X \subseteq I \mid A \subseteq X\}$ ein Filter, der von $A$ **erzeugte Hauptfilter**. Dieser ist ein Ultrafilter genau dann, wenn $A$ nur ein Element besitzt. Interessanter sind Ultrafilter, die keine Hauptfilter sind, sie werden auch **freie** Ultrafilter genannt (während die Hauptultrafilter durch ein Element *fixiert* sind).

2. Hat $I = \{a, b\}$ zwei Elemente $a \neq b$, so gibt es auf $I$ die Filter $\{I\}, \{\{a\}, I\}$ und $\{\{b\}, I\}$; die letzten beiden sind Hauptultrafilter, $\{I\}$ ist zwar Hauptfilter, aber kein Ultrafilter. Allgemeiner ist auf einer endlichen Menge $I$ jeder Filter ein Hauptfilter (erzeugt vom Durchschnitt der endlich-vielen Elemente des Filters).

3. Auf einer unendlichen Menge $M$ ist die Menge

$$\{X \subseteq M \mid X \textit{ co-endlich } (d.h.\ M \setminus X endlich)\}$$

ein Filter auf $M$, der kein Hauptfilter (und auch kein Ultrafilter) ist und welcher im Falle $M = \mathbb{N}$ FRÉCHET-*Filter* genannt wird. Ultrafilter, die den obigen Filter erweitern, können keine endlichen Mengen als Elemente enthalten und sind somit freie Ultrafilter.

Die Existenz von freien Ultrafiltern auf unendlichen Mengen ist nicht trivial; allgemeiner gilt für Mengen $E$ mit der **finite intersection property (endliche-Durchschnitts-Eigenschaft)**

*(fip) Jeder Durchschnitt von endlich-vielen Elementen aus E ist* $\neq \emptyset$:

## 12.3 Ultrafiltersatz, Boolesches Primidealtheorem

(i) *Ist* $E \subseteq \mathcal{P}(I)$ *eine Menge mit der Eigenschaft (fip), so existiert ein Ultrafilter* $U$ *auf* $I$ *mit* $E \subseteq U$. *Insbesondere:*

(ii) *Jeder Filter* $F$ *läßt sich erweitern zu einem Ultrafilter* $U \supseteq F$. $\square$

(Dieses Ergebnis ist äquivalent zum Kompaktheitssatz; am einfachsten beweist man es jedoch mit Hilfe des (stärkeren) ZORNschen Lemmas.)

Ist $U$ ein Ultrafilter auf einer Indexmenge $I$, so können wir die Elemente von $U$ auffassen als Teilmengen von $I$, die “fast alle” Elemente von $I$ enthalten, und die Eigenschaften, die in “fast allen” Strukturen $\mathcal{A}_i$ gelten, werden sich wegen der besonderen “logischen” Eigenschaften von Ultrafiltern auf das entsprechende Ultraprodukt übertragen. Zunächst erklären wir Produkte modulo einem Filter:

## 12.4 Reduziertes Produkt

Es sei $F$ ein Filter auf der Menge $I$ und $(\mathcal{A}_i)_{i \in I}$ eine Familie von $\mathcal{L}$-Strukturen. Für $a, b \in A = \prod_{i \in I} A_i$ setzen wir

$$a \sim_F b :\iff \{i \in I \mid a_i = b_i\} \in F$$

(wir können es lesen als “$a$ ist - bezüglich $F$ - fast überall gleich $b$”) und erhalten eine Äquivalenzrelation auf $A$; mit $a_F$ oder $a/F$ bezeichnen wir die zugehörige Äquivalenzklasse, und für endliche Folgen $\overline{a} = (a^0, \ldots, a^n)$ schreiben wir auch kurz $\overline{a}/F = (a^0/F, \ldots, a^n/F)$.

Das **reduzierte Produkt** $\mathcal{A}/F = \prod_{i \in I} \mathcal{A}_i/F$ wird nun wie folgt definiert:

(R1) $A/F := \{a_F \mid a \in A\}$
ist die Menge der Äquivalenzklassen als Grundbereich,

(R2) $R^{\mathcal{A}_F}(\overline{a}_F) :\Longleftrightarrow \{i \in I \mid R^{\mathcal{A}_i}(\overline{a}_i)\} \in F$
für jedes Relationszeichen $R$ von $\mathcal{L}$,

(R3) $f^{\mathcal{A}_F}(\overline{a}_F) := (f^{\mathcal{A}_i}(\overline{a}_i))_{i \in I}/F = f^{\mathcal{A}}(\overline{a})/F$
für jedes Funktionszeichen $f$ von $\mathcal{L}$,

(R4) $c^{\mathcal{A}/F} := (c^{\mathcal{A}_i})_{i \in I}/F = c^{\mathcal{A}}/F$ für jede Individuenkonstante $c$ von $\mathcal{L}$.

Das bedeutet also, daß im Falle einer 2-stelligen Relation $R$ im reduzierten Produkt

$\overline{a}_F\, R\, \overline{b}_F$ gilt gdw für die Komponenten $a_i R b_i$ in “fast allen” Strukturen $\mathcal{A}_i$ gilt

(wie speziell oben im Falle der Gleichheitsrelation =). (Natürlich muß man nachprüfen, daß die obige Definition korrekt ist, d. h. daß sie nur von den Äquivalenzklassen (und nicht von deren Repräsentanten) abhängig ist.) Sind alle Strukturen $\mathcal{A}_i = \mathcal{A}$, so spricht man von einer **reduzierten Potenz** $\mathcal{A}$, ist der Filter $F$ ein Ultrafilter, von einem **Ultraprodukt** bzw. einer **Ultrapotenz**.

Wir kommen nun zum fundamentalen Satz über Ultraprodukte:

## 12.5 Satz von Łoś

*$I$ sei eine nichtleere Menge und $U$ ein Ultrafilter auf $I$. Ferner sei $(\mathcal{A}_i)_{i \in I}$ eine Familie von $\mathcal{L}$-Strukturen und $\mathcal{A} = \prod_{i \in I} \mathcal{A}_i$ das Produkt.*
*Dann gilt für alle $\mathcal{L}$-Formeln $\varphi$ und alle Belegungen $\overline{a}$:*

$$\mathcal{A}/U \models \varphi[\overline{a}_U] \iff \{i \in I \mid \mathcal{A}_i \models \varphi[\overline{a}_i]\} \in U,$$

d. h. eine Formel gilt im Ultraprodukt gdw sie (für die entsprechenden Komponenten) in fast allen Faktoren gilt.

Zum *Beweis* des Satzes von Łoś führen wir folgende Bezeichnung ein: Für eine $\mathcal{L}$-Formel $\varphi(v_0, \ldots, v_{n-1})$ und ein $n$-Tupel $\overline{a}$ von Elementen aus $A$ heißt

$\|\varphi(\overline{a})\| := \{i \in I \mid \mathcal{A}_i \models \varphi[\overline{a}_i]\}$ die **Boolesche Ausdehnung** von $\varphi(\overline{a})$.

**Lemma**

(i) $\|\neg\varphi(\overline{a})\| = I \setminus \|\varphi(\overline{a})\|$,

(ii) $\|(\varphi \wedge \psi)(\overline{a})\| = \|\varphi(\overline{a})\| \cap \|\psi(\overline{a})\|$,

(iii) $\|(\varphi \vee \psi)(\overline{a})\| = \|\varphi(\overline{a})\| \cup \|\psi(\overline{a})\|$,

(iv) *Für alle (n-1)-Tupel $\overline{a}$ von Elementen aus A und für alle $b \in A$ gilt:*

$$\|\varphi(\overline{a},b)\| \subseteq \|(\exists v_n\varphi)(\overline{a})\|,$$

*und es gibt stets ein $b \in A$ mit*

$$\|\varphi(\overline{a},b)\| = \|(\exists v_n\varphi)(\overline{a})\|.$$

*Beweis:* (i) - (iii) und der erste Teil von (iv) folgen direkt aus der Definition der Booleschen Ausdehnung (und des Wahrheitsprädikates). Um den letzten Teil zu beweisen, beachten wir, daß es zu jedem

$$i \in \|(\exists v_n\varphi)(\overline{a})\| \text{ ein } b_i \in A_i \text{ gibt mit } \mathcal{A}_i \models \varphi[\overline{a}_i, b_i];$$

für jedes andere $j \in I \setminus \|(\exists v_n\varphi)(\overline{a})\|$ wählen wir irgendein $b_j \in A_j$ aus und erhalten damit eine Folge $b = (b_i)_{i \in I}$ mit $b \in A$, für die auch

$$\|(\exists v_n\varphi)(\overline{a})\| \subseteq \|\varphi(\overline{a},b)\|$$

gelten muß. Zusammen mit dem ersten Teil ergibt sich also hier die Gleichheit. □

Dieses Lemma zusammen mit den Eigenschaften (U1)-(U3) eines Ultrafilters ergeben nun leicht einen Beweis des Satzes von Łoś:
Zunächst zeigt man, daß sich (R3) und (R4) der Definition des reduzierten Produktes auf beliebige Terme übertragen lassen:

$$t^{\mathcal{A}/U}(\overline{a}/U) = (t^{\mathcal{A}_i}(\overline{a}_i))_{i \in I}/U = t^{\mathcal{A}}(\overline{a})/U,$$

also gilt (R2) und damit die Behauptung des Satzes - auch für atomare Formeln (also der Form $R(t_1, \dots, t_m)$). Nun benutzen wir Induktion über den Formelaufbau, um sie für beliebige Formeln nachzuweisen.

Sei etwa $\varphi = \neg\psi$ und die Behauptung des Satzes für die Formel $\psi$ bewiesen, d. h. in der neuen Bezeichnungsweise:

$$\begin{array}{rcll} \mathcal{A}/U \models \psi[\overline{a}_U] & \iff & \|\psi(\overline{a})\| \in U, & \text{also auch} \\ \mathcal{A}/U \not\models \psi[\overline{a}_U] & \iff & \|\psi(\overline{a})\| \notin U, & \text{daraus} \\ \mathcal{A}/U \models \neg\psi[\overline{a}_U] & \iff & \|\psi(\overline{a})\| \notin U & \text{nach der Wahrheitsdefinition,} \\ \mathcal{A}/U \models \neg\psi[\overline{a}_U] & \iff & I \setminus \|\psi(\overline{a})\| \in U, & \text{da } U \text{ Ultrafilter und} \\ \mathcal{A}/U \models \neg\psi[\overline{a}_U] & \iff & \|\neg\psi(\overline{a})\| \in U & \text{nach obigem Lemma.} \end{array}$$

Somit haben wir die Behauptung auch für $\varphi = \neg\psi$ bewiesen. Ähnlich argumentiert man im Falle der aussagelogischen Zeichen $\wedge$ und $\vee$. Sei nun schließlich $\varphi = \exists v_n \psi$. Dann gilt

$$\begin{array}{rlll} \mathcal{A}/U \models \exists v_n \psi[\overline{a}_U] & \iff & \mathcal{A}/U \models \psi[\overline{a}_U, b] & \text{für ein } b \in A/U \\ & \iff & \mathcal{A}/U \models \psi[\overline{a}_U, b_U] & \text{für ein } b \in A \\ & \iff & \|\psi(\overline{a}, b)\| \in U & \text{für ein } b \in A \text{ nach Ind.vor.} \\ & \iff & \|\exists v_n \psi(\overline{a})\| \in U & \text{nach obigem Lemma (iv).} \end{array}$$

Damit ist der Satz von Łoś bewiesen. □

**Korollar**

*I sei eine nichtleere Menge, U ein Ultrafilter auf I und $\mathcal{A}$ sei eine $\mathcal{L}$-Struktur. Dann gibt es eine kanonische Einbettung*

$$j : A \longrightarrow A^I/U, \quad \text{definiert durch } j(a) = (a)_{i \in I}$$

*(welche also jedem Element aus A die konstante Folge (a,a,...) zuordnet), und es gilt:*

(i) $j : \mathcal{A} \longrightarrow \mathcal{A}^I/U$ ist eine elementare Einbettung, insbesondere

(ii) $\mathcal{A} \equiv \mathcal{A}^I/U$.

**Beispiele**

1. Für einen Hauptultrafilter sind die obigen Ergebnisse trivial: Wird $U$ durch $\{j\}$ erzeugt, so ist das Ultraprodukt $\mathcal{A} = \prod_{i \in I} \mathcal{A}_i$ isomorph zu $\mathcal{A}_j$, die kanonische Einbettung des Korollars ein Isomorphismus.

2. Ist $\mathcal{A} = \mathcal{N}$ das Standardmodell der natürlichen Zahlen, $U$ ein Ultrafilter auf $I = \mathbb{N}$, so ist die Ultrapotenz $\mathcal{N}^I/U$ elementar-äquivalent zum Standardmodell $\mathcal{N}$, insbesondere auch ein Modell der PEANO-Axiome. Ist $U$ ein freier Ultrafilter, so ist die kanonische Einbettung nicht surjektiv, die Ultrapotenz mithin ein (überabzählbares) Nicht-Standardmodell von PA.

3. Für eine Primzahl $p$ sei $\mathcal{F}_p$ der Primkörper der Charakteristik $p$, $U$ ein freier Ultrafilter auf der Menge $P$ der Primzahlen. Dann ist das Ultraprodukt $\prod_{p \in P} \mathcal{F}_p/U$ ein Körper der Charakteristik 0, denn für die Aussage $\chi_p$, welche ausdrückt, daß ein Körper die Charakteristik $p$ hat, gilt:

$$\{i \in P \mid \mathcal{F}_i \models \chi_p\} = \{p\} \notin U.$$

Wir tragen jetzt den modelltheoretischen Beweis des Kompaktheitssatzes nach (D. SCOTT 1955):

## 12.6 Kompaktheitssatz

*Für jede Theorie* $\mathsf{T}$ *und jeden Satz* $\sigma$ *der Sprache von* $\mathsf{T}$ *gilt:*

(i) $\mathsf{T} \models \sigma \Longrightarrow$ *es gibt ein endliches* $\mathsf{T}_0 \subseteq \mathsf{T}$ *mit* $\mathsf{T} \models \sigma$,

(ii) *jedes endliche* $\mathsf{T}_0 \subseteq \mathsf{T}$ *hat ein Modell* $\Longrightarrow \mathsf{T}$ *hat ein Modell.*

*Beweis* von (ii): $\Sigma$ sei die Menge der Axiome von $\mathsf{T}$, wobei wir annehmen können, daß alle Formeln in $\Sigma$ Sätze sind (sonst ersetze man sie jeweils durch ihren universellen Abschluß). Nach Voraussetzung gibt es für jedes endliche $\Delta \subseteq \Sigma$ ein Modell $\mathcal{A}_\Delta \models \Delta$; wir wählen also als Indexmenge

$$I := \{\Delta \mid \Delta \subseteq \Sigma,\ \Delta \ endlich\}$$

und werden ein Modell von $\Sigma$ als ein Ultraprodukt finden für einen geeigneten Ultrafilter auf $I$. Dazu definieren wir für jeden Satz $\sigma \in \Sigma$ die Menge

$$\hat{\sigma} := \{\Delta \in I \mid \sigma \in \Delta\} \subseteq I \quad \text{und setzen} \quad E := \{\hat{\sigma} \mid \sigma \in \Sigma\} \subseteq \mathcal{P}(I).$$

$E$ hat die Eigenschaft *(fip)* des Ultrafiltersatzes 12.3, da

$$\hat{\sigma}_1, \dots, \hat{\sigma}_n \in E \Longrightarrow \{\sigma_1, \dots, \sigma_n\} \in \hat{\sigma}_1 \cap \dots \cap \hat{\sigma}_n,$$

also existiert ein Ultrafilter $U$ auf $I$ mit $E \subseteq U$. Wir behaupten nun, daß

$$\mathcal{A} = \prod_{\Delta \in I} \mathcal{A}_\Delta / U \models \Sigma :$$

Sei $\sigma \in \Sigma$. Dann gilt für alle $\Delta \in I$:

$$\Delta \in \hat{\sigma} \Longrightarrow \sigma \in \Delta \Longrightarrow \mathcal{A}_\Delta \models \sigma.$$

Somit

$$\begin{aligned} \hat{\sigma} \subseteq \{\Delta \in I \mid \mathcal{A}_\Delta \models \sigma\}, &\quad \text{und wegen} \\ \hat{\sigma} \in U &\quad \text{ist auch die Obermenge} \\ \{\Delta \in I \mid \mathcal{A}_\Delta \models \sigma\} \in U. & \end{aligned}$$

Letzteres besagt aber nach dem Satz von Łoś: $\mathcal{A} \models \sigma$.

(i) folgt aus (ii), indem man die Tatsache benutzt, daß für jeden Satz $\sigma$ gilt:

$$\mathsf{T} \models \sigma \iff \mathsf{T} \cup \{\neg\sigma\} \text{ hat ein Modell.}$$

□

**Weitere Folgerungen aus dem Kompaktheitssatz:**

1. Ist $\mathcal{K}$ eine endlich-axiomatisierbare Klasse von $\mathcal{L}$-Strukturen (also eine EC-Klasse), so gibt es zu *jedem* Axiomensystem $\Sigma$ für $\mathcal{K}$ eine endliche Teilmenge von $\Sigma$, welche $\mathcal{K}$ axiomatisiert.

2. $\mathcal{K}$ ist eine *EC*-Klasse genau dann, wenn $\mathcal{K}$ und das Komplement von $\mathcal{K}$ $EC_\Delta$-Klassen sind.

3. Eine Klasse $\mathcal{K}$ von $\mathcal{L}$-Strukturen ist eine $EC_\Delta$-Klasse gdw $\mathcal{K}$ abgeschlossen ist unter Ultraprodukten und elementarer Äquivalenz.

# Kapitel 13

# Modellvollständigkeit

**Definition**

(1) Eine Theorie $\mathsf{T}$ heißt **modellvollständig** gdw

$$\mathsf{T} \cup \mathsf{D}(\mathcal{A}) \text{ ist vollständig (bzgl. der Sprache } \mathcal{L}_A) \text{ für jedes } \mathcal{A} \models \mathsf{T}.$$

Mit Hilfe des Diagrammlemmas 10.2 ergibt sich, daß eine Theorie $\mathsf{T}$ modellvollständig ist gdw

$$\text{für alle } \mathcal{A}, \mathcal{B} \models \mathsf{T} \text{ gilt: } \mathcal{A} \subseteq \mathcal{B} \Longrightarrow \mathcal{A} \preccurlyeq \mathcal{B}.$$

Insbesondere folgt aus dem Kettenlemma 11.1 und dem Satz 11.2 von CHANG-ŁOŚ-SZUSKO, daß jede modellvollständige Theorie induktiv ist, also ein Axiomensystem aus $\forall\exists$-Sätzen besitzt.

(2) Ein Modell $\mathcal{P} \models \mathsf{T}$ einer Theorie $\mathsf{T}$ heißt **Primmodell** von $\mathsf{T}$ gdw es sich in jedes Modell von $\mathsf{T}$ einbetten läßt.

**Beispiele und Bemerkungen**

1. $\mathcal{A}_A$ ist (wegen des Diagrammlemmas 10.2) stets ein Primmodell von $\mathsf{D}(\mathcal{A})$.

2. Ist die Theorie $\mathsf{T}$ modellvollständig und besitzt $\mathsf{T}$ ein Primmodell, so ist $\mathsf{T}$ vollständig.

3. In der Algebra wird der Begriff des *Primkörpers* eingeführt: dieser ist ein Primmodell in obigem Sinne, und zwar ist der Körper $\mathbb{Q}$ der rationalen Zahlen ein Primmodell der Theorie der Körper der Charakteristik 0, während

$\mathbb{Z}/p$, der Restklassenkörper modulo $p$, das Primmodell der Theorie der Körper der Charakteristik $p$ ist ($p$ eine Primzahl). Die Theorie der Körper $\mathsf{T}_{\mathsf{F}}$ allein besitzt offenbar keinen Primkörper, sie ist weder vollständig noch modellvollständig.

4. Der Körper der reellen algebraischen Zahlen ist Primkörper der Theorie $\mathsf{T}_{\mathsf{RCF}}$ der reell-abgeschlossenen Körper, diese Theorie ist modellvollständig und somit vollständig.

5. Der Körper der algebraischen Zahlen ist Primkörper der Theorie der algebraisch-abgeschlossenen Körper der Charakteristik 0. Die Theorie $\mathsf{T}_{\mathsf{ACF}}$ der algebraisch-abgeschlossenen Körper (ohne Festlegung der Charakteristik) ist zwar modellvollständig, aber nicht vollständig.

Die obige Charakterisierung der modellvollständigen Theorien deutet darauf hin, daß man in solchen Theorien Eigenschaften (fast) quantorenfrei ausdrücken kann:

**Definition**

$\mathsf{T}$ sei eine Theorie der Sprache $\mathcal{L}$, $\mathcal{A}$ eine $\mathcal{L}$-Struktur.

1. $\mathsf{T}$ läßt (schwache) **Quantorenelimination** zu gdw zu jeder $\mathcal{L}$-Formel $\varphi$ eine quantorenfreie (bzw. $\exists$-) Formel $\psi$ existiert mit

$$\mathsf{T} \models \varphi \leftrightarrow \psi.$$

2. $\mathcal{A} \preccurlyeq_1 \mathcal{B} :\Longleftrightarrow \mathcal{A} \subseteq \mathcal{B}$ und für alle $\exists$-Formeln $\psi(v_1,\dots,v_n)$ und alle $a_1,\dots a_n \in A$ gilt:

$$\mathcal{A} \models \varphi[a_1,\dots a_n] \iff \mathcal{B} \models \varphi[a_1,\dots a_n].$$

3. $\mathcal{A}$ heißt $\mathsf{T}$- e.c. ($\mathsf{T}$**-existentiell-abgeschlossen**) gdw für alle $\mathcal{L}$-Strukturen $\mathcal{B}$ gilt:

$$\mathcal{A} \subseteq \mathcal{B} \models \mathsf{T} \Longrightarrow \mathcal{A} \preccurlyeq_1 \mathcal{B}.$$

4. $\mathsf{E}_{\mathsf{T}} := \{\mathcal{A} \mid \mathcal{A}\ \mathsf{T}\text{- e.c.}\}$.

## 13.1 Robinsonscher Test

*Für eine Theorie* $\mathsf{T}$ *sind folgende Aussagen äquivalent:*

(i) $\mathsf{T}$ *ist modellvollständig,*

(ii) $\mathsf{T}$ *läßt schwache Quantorenelimination zu,*

(iii) *jede* $\exists$*-Formel* $\varphi$ *ist in* $\mathsf{T}$ *äquivalent zu einer* $\forall$*-Formel* $\delta$*,*

(iv) *für alle* $\mathcal{A},\mathcal{B} \models \mathsf{T}$ *gilt:* $\mathcal{A} \subseteq \mathcal{B} \Longrightarrow \mathcal{A} \preccurlyeq_1 \mathcal{B}$*,*

(v) *für alle* $\mathcal{A},\mathcal{B} \models \mathsf{T}$ *gilt: Ist* $f : \mathcal{A} \to \mathcal{B}$ *eine Einbettung, so ist f eine elementare Einbettung.*

(vi) $Mod(\mathsf{T}) = \mathsf{E}_{\mathsf{T}}$.

Hierin ist (iv) das eigentliche Test-Kriterium, das insbesondere für Anwendungen auf algebraische Theorien noch etwas abgeschwächt werden kann.[1]

Um ein Kriterium für die vollständige Elimination von Quantoren zu erhalten, muß man die Eigenschaft der Modellvollständigkeit verstärken:

### Definition

Eine Theorie $\mathsf{T}$ heißt **substrukturvollständig** gdw

$$\mathsf{T} \cup \mathsf{D}(\mathcal{A}) \text{ ist vollständig für jedes } \mathcal{A} \models \mathsf{T}_\forall,$$

d. h. $\mathsf{T} \cup \mathsf{D}(\mathcal{A})$ ist vollständig für jede Substruktur $\mathcal{A}$ eines Modells $\mathcal{A} \subseteq \mathcal{B} \models \mathsf{T}$.

### Satz

*Für eine Theorie* $\mathsf{T}$ *sind folgende Aussagen äquivalent:*

(i) $\mathsf{T}$ *ist substrukturvollständig,*

(ii) $\mathsf{T}$ *läßt Quantorenelimination zu.*

*Insbesondere gilt für eine* $\forall$*-Theorie* $\mathsf{T}$*:*
$\mathsf{T}$ *läßt Quantorenelimination zu gdw* $\mathsf{T}$ *modellvollständig ist.*

[1] S. etwa das Buch von RAUTENBERG, W.: Einführung in die mathematische Logik, vieweg 2002, Kap. 5.5

**Beispiele**

1. Die Theorie $\mathsf{T}_{\mathsf{ACF}}$ der algebraisch-abgeschlossenen Körper ist modellvollständig, legt man auch die Charakteristik fest, so ist sie vollständig.

2. Die Theorie $\mathsf{T}_{\mathsf{RCF}}$ der reell-abgeschlossenen Körper ist modellvollständig und vollständig. Sie stimmt überein mit der Theorie des angeordneten Körpers der reellen Zahlen.

3. Die Theorie der Vektorräume über den rationalen Zahlen ist modellvollständig (eine $\exists$-Aussage über eine Konjunktion von Basisformeln besagt in diesem Falle die Lösbarkeit eines Systems von linearen Gleichungen und Ungleichungen).

4. Die Theorie der Gleichheit (also ohne nicht-logischen Axiome!) ist weder vollständig noch modellvollständig. Sie erlaubt auch keine Quantorenelimination, allerdings ist in dieser Theorie jede Formel äquivalent zu einer aussagenlogisch aus Anzahlformeln und den Formeln $v_i = v_j$ zusammengesetzten Formel.

5. Die Theorie $\mathsf{DLO} + \mathsf{OE}$ der dichten linearen Ordnungen ohne Endpunkte ist vollständig, modellvollständig und erlaubt Quantorenelimination.

6. Die Theorie $\mathsf{Th}(\mathbb{N}, <)$ ist vollständig, aber nicht modellvollständig.

Theorien mit Quantorenelimination sind meistens *entscheidbar*, d. h. es gibt ein *effektives Verfahren* (dieser Ausdruck wird später präzisiert werden), welches für jeden Satz der Sprache von $\mathsf{T}$ angibt, ob dieser Satz in $\mathsf{T}$ gilt oder nicht. So sind die Theorien

- $\mathsf{DLO} + \mathsf{OE}$,
- die Theorie des (angeordneten) Körpers der reellen Zahlen (= Theorie der reell-abgeschlossenen Körper) und
- die Theorie des Körpers der komplexen Zahlen (= Theorie der algebraisch-abgeschlossenen Körper der Charakteristik 0)

vollständig und entscheidbar. Dagegen ist dagegen die Theorie der natürlichen Zahlen $\mathsf{Th}(\mathbb{N}, +, \cdot, 0, 1)$ unentscheidbar. Näher gehen wir hierauf in den abschließenden Kapiteln ein.

# Teil IV

# Unvollständigkeit und Unentscheidbarkeit

## Hans Magnus Enzensberger: *Hommage à Gödel*

Münchhausens Theorem, Pferd, Sumpf und Schopf,
ist bezaubernd, aber vergiß nicht:
Münchhausen war ein Lügner.

Gödels Theorem wirkt auf den ersten Blick
etwas unscheinbar, doch bedenk:
Gödel hat recht.

»In jedem genügend reichhaltigen System // lassen sich Sätze formulieren,
die innerhalb des Systems // weder beweis- noch widerlegbar sind,
es sei denn das System // wäre selber inkonsistent.«

Du kannst deine eigene Sprache // in deiner eigenen Sprache beschreiben:
aber nicht ganz.
Du kannst dein eignes Gehirn // mit deinem eignen Gehirn erforschen:
aber nicht ganz.
Usw.

Um sich zu rechtfertigen // muß jedes denkbare System
sich transzendieren, // d.h. zerstören.

»Genügend reichhaltig« oder nicht:
Widerspruchsfreiheit
ist eine Mangelerscheinung
oder ein Widerspruch.

(Gewißheit = Inkonsistenz.)

Jeder denkbare Reiter, // also auch Münchhausen,
also auch du bist ein Subsystem // eines genügend reichhaltigen Sumpfes.

Und ein Subsystem dieses Subsystems // ist der eigene Schopf,
dieses Hebezeug // für Reformisten und Lügner.

In jedem genügend reichhaltigen System,
also auch in diesem Sumpf hier,
lassen sich Sätze formulieren,
die innerhalb des Systems
weder beweis- noch widerlegbar sind.

Diese Sätze nimm in die Hand
und zieh!

(Zitiert nach: Hans Magnus Enzensberger: *Die Elixiere der Wissenschaft*.
Frankfurt a.M.: Suhrkamp 2002)

In diesem Teil gehen wir aus von den natürlichen Zahlen als algebraische Struktur mit den Operationen der Addition und Multiplikation und zusätzlicher Ordnung. Als Axiomensystem für die elementaren Eigenschaften dieser Struktur haben wir bereits das Axiomensystem von G. PEANO erwähnt und gesehen, daß es nicht ausdrucksstark genug ist, um die Struktur der natürlichen Zahlen bis auf Isomorphie festzulegen. Tatsächlich ist es sogar nach einem berühmten Satz von K. GÖDEL *unvollständig*, und wir werden sehen, daß dieses Problem sich auch durch die Hinzunahme weiterer Axiome nicht beseitigen läßt.

Dazu werden wir mit den Methoden der mathematischen Logik Prädikate und Funktionen natürlicher Zahlen nach ihrer *Komplexität* klassifizieren, insbesondere den Begriff der *Berechenbarkeit* und der *Entscheidbarkeit* präzisieren, um damit Möglichkeiten und Grenzen der Beweisbarkeit für mathematischer Theorien, vor allem für die Zahlentheorie und auch die Logik selbst aufzuzeigen.

Eine ausführlichere Darstellung der Theorie der Berechenbarkeit findet man im Skriptum

*K. Ambos-Spies: Theoretische Informatik*
(http://www.math.uni-heidelberg.de/logic/skripten.html).

# Kapitel 14

# Berechenbare Funktionen

## 14.1 Turing-Maschinen

Eine *TURING-Maschine M* besitzt als *Speicher* ein nach beiden Seiten unbeschränktes *Band*, das in unendlich viele *Felder* eingeteilt ist. Jedes Feld kann einen Buchstaben aus dem endlichen *Alphabet* $s_1, \ldots, s_n$ der Maschine $M$ aufnehmen. Dabei sind zu jedem Zeitpunkt nur endlich viele Felder belegt, wobei wir vereinbaren können, daß die leeren Feldern durch ein zusätzliches Symbol $s_0$ gekennzeichnet sind. Der Zugriff auf das Speicherband erfolgt durch einen kombinierten *Lese- und Schreibkopf* von der Größe eines Feldes. Das Feld, auf das der Kopf zeigt, heißt das *Arbeitsfeld*. Dieses Feld kann eingelesen und neu beschriftet werden. Weiterhin kann sich der Kopf um ein Feld nach links oder rechts bewegen. Durch Wiederholung dieser elementaren Operationen kann also jede Stelle des Bandes besucht, die dort stehende Information gelesen und gegebenenfalls ersetzt oder ergänzt werden. Außerdem ist $M$ zu jedem gegebenen Zeitpunkt in einem von endlich-vielen *Zuständen* $q_1, \ldots, q_r$, die Einfluß auf das Verhalten von $M$ haben.

Es gibt somit 3 Typen von Operationen, genannt *Anweisungen*, die $M$ ausführen kann, wenn $M$ im Zustand $q_i$ über einem Feld steht, in dem das Zeichen $s_j$ eingetragen ist, und die wie folgt bezeichnet werden:

(a) $q_i\, s_j\, s_k\, q_l$: lösche $s_j$ und schreibe dafür $s_k$,

(b) $q_i\, s_j\, R\, q_l$: bewege den Kopf ein Feld weiter nach rechts,

(c) $q_i\, s_j\, L\, q_l$: bewege den Kopf ein Feld weiter nach links.

Anschließend gehe $M$ in allen Fällen in den Zustand $q_l$ über.

Ein *Programm* für $M$ ist eine endliche Liste von Anweisungen der obigen Formen (a) - (c), welches konsistent ist, d. h. zu einem gegebenen Paar $q_i, s_j$ (für die Ausgangssituation) gibt es in der Liste höchstens eine Anweisung, die hiermit beginnt.

**Beispiel:**

Das Alphabet von $M$ bestehe aus den Symbolen $0, 1$ (und $x$ für ein leeres Feld), mögliche Zustände seien $q_1, q_2$. Das Programm $P$ bestehe aus den folgenden 4 Anweisungen:

$$q_1\,0\,R\,q_1$$
$$q_1\,1\,0\,q_2$$
$$q_2\,0\,R\,q_2$$
$$q_2\,1\,R\,q_1$$

Anfangs stehe $M$ im Zustand $q_1$ über dem ersten Feld des Bandes, das von hier ab weitere $n$ aufeinander folgende Felder mit dem Symbol 1 enthält, während die übrigen Felder leer sind:

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | ↓ | | | | | | | | | |
| … | x | x | x | 1 | 1 | 1 | 1 | 1 | 1 | x | x | x | … |

Folgt $M$ den obigen Anweisungen, so kann es diese ausführen, bis der folgende Endzustand erreicht ist:

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | ↓ | | | |
| … | x | x | x | 0 | 1 | 0 | 1 | 0 | 1 | x | x | x | … |

Wir wollen nun erklären, wie eine T-Maschine mit vorgegebenen Programm eine Funktion berechnet. Da es vorkommen kann, daß Rechnungsprogramme bei Anwendungen auf bestimmte Anfangswerte möglicherweise nicht stoppen (weil sie z. B. in eine Schleife geraten), werden wir jetzt statt *totaler* Funktionen $f : \mathbb{N}^k \to \mathbb{N}$ auch *partielle* Funktionen zulassen:

**Definition**

$f$ ist eine **partielle** $k$-stellige Funktion auf den natürlichen Zahlen gdw

$$f : D \to \mathbb{N} \text{ für eine Teilmenge } D \subseteq \mathbb{N}^k.$$

In diesem Falle bedeute

$$
\begin{aligned}
f(\vec{x})\downarrow: &\iff f \text{ ist an der Stelle } \vec{x} \text{ definiert}\\
f(\vec{x})\uparrow: &\iff f \text{ ist an der Stelle } \vec{x} \text{ nicht definiert}\\
f(\vec{x}) \simeq g(\vec{x}): &\iff (f(\vec{x})\downarrow \iff g(\vec{x})\downarrow) \text{ und } (f(\vec{x})\downarrow \Longrightarrow f(\vec{x}) = g(\vec{x})).
\end{aligned}
$$

**TURING-berechenbare Funktionen**

Zur Darstellung von Zahlen, von denen eine Rechnung ausgeht, wählen wir $s_1 = 1$ und benutzen es als Marke, indem $n+1$ aufeinander folgende Felder mit Inhalt 1 (umgeben von Leerfeldern am Anfang und am Ende) die Zahl $n$ darstellen. Eine partielle Funktion $f : D \to \mathbb{N}$ wird von $M$ **berechnet**, wenn $M$, beginnend im Anfangszustand $q_1$ über dem ersten Feld, welches die Zahl $n$ darstellt, genau dann über dem Band in einem Endzustand stoppt, wenn $n \in D$ ist, und in diesem Fall das Symbol 1 genau $f(n)$-mal auf dem Band im Endzustand eingetragen ist (in obigem Beispiel ist $f(5) = 3$). Eine ähnliche Definition legt man im Falle mehrstelliger Funktionen fest.

Eine partielle Funktion $f$ heißt **TURING-berechenbar** genau dann, wenn es eine T-Maschine $M$ gibt mit einem Programm, welches $f$ berechnet.

## 14.2 URM-berechenbare Funktionen

SHEPERDSON-STURGIS haben 1963 eine Variante eines Maschinenmodells vorgeschlagen: Eine *unbeschränkte Register-Maschine* (URM) besitzt eine unendliche Anzahl von *Registern* $R_1, \dots, R_l$, die jeweils eine Zahl enthalten; $r_n$ sei die Zahl im Register $R_n$. Diese Zahlen werden von einer URM gemäß folgenden Typen von *Anweisungen* abgeändert:

| | | | | |
|---|---|---|---|---|
| $O(n)$: | $0 \to R_n$ | bzw. | $r_n := 0$ | ersetze $r_n$ durch 0 |
| $S(n)$: | $r_n + 1 \to R_n$ | bzw. | $r_n := r_n + 1$ | addiere 1 zu $r_n$ |
| $T(m,n)$: | $r_m \to R_n$ | bzw. | $r_n := r_m$ | ersetze $r_n$ durch $r_m$ |
| $J(m,n,q)$: | falls $r_n = r_m$, | so gehe zur $q$-ten Anweisung über, sonst zur nächsten. | | |

Ein *Programm* ist nun wieder eine nummerierte endliche Folge derartiger Anweisungen. Soll eine URM mit einem Programm $P = (I_1, \dots, I_k)$ eine Berechnung durchführen, so muß ein Anfangszustand der Registerinhalte gegeben sein, in welchem die URM mit der ersten Anweisung beginnt und worauf sie dann den weiteren Anweisungen der Liste folgt.

**Beispiel:**

$$\begin{array}{ll} I_1: & J(1,2,6) \\ I_2: & S(2) \\ I_3: & S(3) \\ I_4: & J(1,2,6) \\ I_5: & J(1,1,2) \\ I_6: & T(3,1) \end{array}$$

Beginnt die URM mit den Zahlen 9 im 1. und 5 im 2. Register (und 0 in allen anderen), so erzeugt sie nach einigen Schritten die Zahl 4 im 3. Register und landet dann bei der Anweisung $I_6$, welche die Zahl 4 in das 1. Register überträgt; da keine weitere Anweisung folgt, stoppt hier das Programm.

Es sei nun $f$ eine partielle $n$-stellige Funktion auf den natürlichen Zahlen, $P$ ein Programm und $a_1, \ldots, a_n, b$ seien vorgegebene Zahlen. Dann *konvergiert* das mit $a_1, \ldots, a_n$ beginnende Programm *gegen die Zahl b*, formal: $P(a_1, \ldots, a_n) \downarrow b$, gdw es nach endlich-vielen Schritten stoppt und die Zahl $b$ im 1. Register steht.

*P berechnet f* gdw für alle $a_1, \ldots, a_n, b$ gilt:

$$P(a_1, \ldots, a_n) \downarrow b \text{ gdw } f(a_1, \ldots, a_n) \text{ ist definiert und} = b.$$

$f$ ist **URM-berechenbar** gdw es ein Programm $P$ gibt, welches $f$ berechnet.

## 14.3 Churchsche These

Es gibt weitere Ansätze, den Begriff der *berechenbaren Funktion* zu präzisieren:

- GÖDEL-HERBRAND-KLEENE (1936): mittels eines *Gleichungskalküls* definierte *allgemein-rekursive* Funktionen
- CHURCH (1936): $\lambda$-definierbare Funktionen
- GÖDEL-KLEENE (1936): $\mu$-rekursive Funktionen (s.u.)
- POST (1943): berechenbare Funktionen mittels *kanonischer Deduktionssysteme*
- MARKOV (1951): berechenbare Funktionen mittels *Algorithmen* über endlichen Alphabeten

Alle diese Definitionen haben sich als formal äquivalent herausgestellt, außerdem ist bisher jede einzelne der (im intuitiven Sinne) berechenbaren Funktionen auch berechenbar in dem obigen formalen Sinne, was Anlaß zu folgender Annahme gegeben hat:

**Churchsche These:** Die (intuitiv) berechenbaren Funktionen sind genau die (in einem der obigen Sinne) formal berechenbaren Funktionen.

Ohne ausführliche Beweise ergänzen wir einige wichtige

## 14.4 Aufzählbarkeitssätze

- Da sich alle Programme effektiv aufzählen lassen, gibt es eine effektive Aufzählung $\varphi_0, \varphi_1, \ldots, \varphi_n, \ldots$ aller einstelligen berechenbaren (partiellen) Funktionen. Mit einem Diagonalargument erhält man hieraus:
- es gibt eine totale Funktion $f : \mathbb{N} \to \mathbb{N}$, die *nicht* berechenbar ist:
  $$f(n) = \begin{cases} \varphi_n(n) + 1 & \text{falls } \varphi_n(n) \text{ definiert ist,} \\ 0 & \text{sonst.} \end{cases}$$
- Somit gibt es zwar eine Abzählung aller *totalen* berechenbaren Funktionen, aber diese Abzählung kann selbst nicht berechenbar sein.
- Auch die folgenden Prädikate sind *nicht entscheidbar* (d. h. ihre charakteristischen Funktionen sind nicht berechenbar):
  $\varphi_n(n)$ ist definiert, $\varphi_n$ ist totale Funktion, $\varphi_n$ ist die konstante Funktion 0, $\varphi_n = \varphi_m$, das $n$-te Programm mit Eingabe $n$ (allgemeiner mit Eingabe $m$) stoppt nach endlich-vielen Schritten (das **Halteproblem**).
- Existenz **universaler** Funktionen (bzw. Programme): Zu jeder berechenbaren 2-stelligen Funktion $f$ gibt es eine totale berechenbare Funktion $k$ mit $f(n,m) = \varphi_{k(n)}(m)$; die 2-stellige Funktion $\varphi(n,m) = \varphi_n(m)$ ist berechenbar.

## 14.5 Primitiv-rekursive Funktionen

Zu den intuitiv berechenbaren Funktionen gehören sicher die folgenden

**Anfangsfunktionen**

| | | |
|---|---|---|
| R1 | $0(x) = 0$ | *Nullfunktion* |
| R2 | $S(x) = x+1$ | *Nachfolgerfunktion* |
| R3 | $P_i^n(x_1,\ldots,x_n) = x_i$ | *i-te Projektion* |

Durch Einsetzen von Funktionen in bereits vorgegebene Funktionen erhalten wir eine neue Funktion durch

**Substitution**

R4 $f(\vec{x}) \simeq h(g_1(\vec{x}),\ldots,g_k(\vec{x}))$

Wichtig ist nun das folgende Prinzip - nach diesem Muster werden nämlich die arithmetischen Operationen Addition, Multiplikation und Potenz gebildet:

**Primitive Rekursion**

R5
$$\begin{aligned} f(\vec{x},0) &\simeq g(\vec{x}) \\ f(\vec{x},y+1) &\simeq h(\vec{x},y,f(\vec{x},y)) \end{aligned}$$

Die Klasse **PRIM** aller **primitiv-rekursiven** (abgekürzt: **p.r.**) Funktionen ist die kleinste Klasse von Funktionen, die die Anfangsfunktionen enthält und abgeschlossen ist unter Substitution und primitiver Rekursion. Diese Funktionen sind insbesondere totale Funktionen.

**Beispiele primitiv-rekursiver Funktionen**

*Die folgenden Funktionen sind primitiv-rekursiv:*

| | |
|---|---|
| $x+y$ | *Summe* |
| $x \cdot y$ | *Produkt* |
| $x^y$ | *Potenz* |
| $\max(x,y)$ | *Maximum* |
| $\min(x,y)$ | *Minimum* |
| $\lvert x-y \rvert$ | *absolute Differenz* |
| $x \dot{-} y = \begin{cases} x-y & \text{falls } x \geq y, \\ 0 & \text{sonst} \end{cases}$ | *Differenz auf* $\mathbb{N}$ |

$$sg(x) = \begin{cases} 0 & \text{falls } x = 0, \\ 1 & \text{falls } x > 0 \end{cases}$$ *Signum, Vorzeichen*

$$\overline{sg}(x) = \begin{cases} 1 & \text{falls } x = 0, \\ 0 & \text{falls } x > 0 \end{cases}$$ *Antisignum, negiertes Vorzeichen*

*Beweis:* Addition, Multiplikation und die Potenz sind offenbar primitiv-rekursiv. Für die übrigen Fälle benötigt zeigt man zuerst, daß die Vorgängerfunktion $x \dot{-} 1$ primitiv-rekursiv ist, und zwar wegen

$$\begin{array}{lcl} 0 \dot{-} 1 & = & 0 \\ (x+1) \dot{-} 1 & = & x \end{array}$$

und definiert dann $\dot{-}$ durch primitive Rekursion

$$\begin{array}{lcl} x \dot{-} 0 & = & x \\ x \dot{-} (y+1) & = & (x \dot{-} y) \dot{-} 1 \end{array}$$

Die übrigen Funktionen kann man direkt mittels $+$ und $\dot{-}$ definieren:

$$\begin{array}{lcl} |x-y| & = & (x \dot{-} y) + (y \dot{-} x) \\ \max(x,y) & = & x + (y \dot{-} x) \\ \min(x,y) & = & \max(x,y) \dot{-} |x-y| \\ \overline{sg}(x) & = & 1 \dot{-} x \\ sg(x) & = & 1 \dot{-} \overline{sg}(x) \end{array}$$

□

**Abschlußeigenschaften**

*Primitiv-rekursive Funktionen sind abgeschlossen unter*

(i) **Fallunterscheidung**:

*Sind $g, f_0, f_1, \ldots, f_k$ primitiv-rekursiv, so auch f mit*

$$f(\vec{x}) = \begin{cases} f_0(\vec{x}) & falls \quad g(\vec{x}) = 0, \\ f_1\vec{x}) & falls \quad g(\vec{x}) = 1, \\ \vdots & \vdots \\ f_k(\vec{x}) & falls \quad g(\vec{x}) \geq k. \end{cases}$$

(ii) **beschränktem $\mu$-Operator**:

*Ist g primitiv-rekursiv, so auch f mit*

$$f(\vec{x},z) = \mu y < z\,(g(\vec{x},y) = 0),$$

wobei

$$\mu y < z\,R(\vec{x},y) = \begin{cases} \textit{das kleinste } y < z \textit{ mit } R(\vec{x},y) & \textit{falls ein solches existiert} \\ z & \textit{sonst}. \end{cases}$$

## 14.6 Rekursive und partiell-rekursive Funktionen

Ein Beispiel einer berechenbaren, aber nicht p.r. Funktion ist die

**Die Ackermannsche Funktion:**

Diese ist eine 2-stellige Funktion, die durch die folgenden Gleichungen rekursiv bestimmt ist:

$$\begin{aligned} A(0,y) &= y+1 \\ A(x+1,0) &= A(x,1) \\ A(x+1,y+1) &= A(x,A(x+1,y)) \end{aligned}$$

Es handelt sich hier um eine doppelte Rekursion, trotzdem kann man leicht nachprüfen, daß sich jeder Wert $A(x,y)$ von endlich-vielen "früheren" Werten $A(u,v)$ mit $u < x$ oder $u = x \wedge v < y$ bestimmen läßt. Schreiben wir $F_n(m)$ für $A(n,m)$, so erhalten wir:

$$\begin{aligned} F_0(m) &= m+1 \\ F_1(m) &= m+2 \\ F_2(m) &\approx 2m \\ F_3(m) &\approx 2^m \\ F_4(m) &\approx 2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}\ (m-mal) \end{aligned}$$

Das bedeutet also, daß die ACKERMANNsche Funktion die Grundrechenarten iteriert und insbesondere in der ersten Stelle (die die Anzahl der Iterationen angibt) sehr schnell wächst. Tatsächlich majorisiert sie jede p.r. Funktion: Ist $f$ eine $k$-stellige p.r. Funktion, so gibt es eine Zahl $n$ mit

$$f(x_1,\dots,x_k) \le F_n(\max(x_1,\dots,x_k))$$

für alle $x_1,\dots,x_k$. Somit kann die 2-stellige Funktion $A$ *nicht* p.r. sein (obwohl andererseits alle $F_n$ p.r. sind!). Daß die ACKERMANNsche Funktion trotzdem berechenbar ist, ersieht man daraus, daß sie aus einer p.r. Funktion erhalten werden kann mittels

### Minimalisierung

R6 $f(\vec{x}) \simeq \mu y\,(g(\vec{x},y) \simeq 0)$,

d. h.

$$f(\vec{x}) = \begin{cases} \textit{das kleinste } y \textit{ so daß} & \\ \quad (i)\ g(\vec{x},z) \textit{ definiert ist für alle } z \le y \textit{ und} & \\ \quad (ii)\ g(\vec{x},y) = 0, & \text{falls ein derartiges } y \text{ existiert,} \\ \textit{undefiniert} & \text{sonst.} \end{cases}$$

Dieses Prinzip kann von totalen zu partiellen Funktionen führen (z. B. für $g(x,y) = |x - y^2|$). Die Klasse **R** (oder $\mathcal{C}$) aller **rekursiven** Funktionen ist die kleinste Klasse von Funktionen, die die Anfangsfunktionen enthält und abgeschlossen ist unter Substitution, primitiver Rekursion und Minimalisierung. Man spricht mitunter auch von *allgemein-rekursiven*, *$\mu$-rekursiven* Funktionen oder (auf Grund der CHURCHschen These) von den berechenbaren Funktionen ($\mathcal{C}$ = *computable functions*). Von *rekursiven* Funktionen werden wir vorwiegend im Falle totaler berechenbarer Funktionen sprechen.

Jede p.r. Funktion ist also auch eine rekursive Funktion (während die ACKERMANNsche Funktion ein Beispiel für eine rekursive, aber nicht p.r. Funktion ist), und die Abschlußeigenschaften der p.r. Funktionen gelten (geeignet modifiziert für partielle Funktionen) auch für rekursive Funktionen.

### Definition

Eine zahlentheoretische Relation $R \subseteq \mathbb{N}^k$ heißt **primitiv-rekursiv** bzw. **rekursiv** (oder auch: **entscheidbar**) gdw ihre charakteristische Funktion $c_R$ primitiv-rekursiv bzw. rekursiv ist, wobei

$$c_R(\vec{x}) = \begin{cases} 1 & \text{falls } R(\vec{x}), \\ 0 & \text{sonst.} \end{cases}$$

Somit sind z. B. die Prädikate $x = y, x \neq y, x \leq y, x < y$ primitiv-rekursiv.

## 14.7 Partielle Entscheidbarkeit

Eine Relation $R \subseteq \mathbb{N}^k$ heißt **partiell-** (oder **positiv-**) **entscheidbar** bzw. **r.e.** gdw. sie Definitionsbereich einer berechenbaren Funktion ist:

$$R\ r.e. :\Longleftrightarrow R = \{\vec{x} \mid f(\vec{x}) \downarrow\} \text{ für eine berechenbare Funktion } f.$$

In diesem Fall kann man für $f$ auch die **partielle charakteristische** Funktion von $R$ wählen:

$$R\ r.e. \iff pc_R \textit{ ist rekursiv},$$

wobei $$pc_R(\vec{x}) = \begin{cases} 1 & \text{falls } \vec{x} \in R, \\ \textit{undefiniert} & \text{sonst.} \end{cases}$$

Während es also für eine *entscheidbare* Relation $R$ ein effektives Verfahren gibt, welches die Frage beantwortet, ob $R$ zutrifft oder nicht, kann man für eine *partiell-entscheidbare* Relation somit nur den Fall bestätigen, daß die Relation zutrifft. Die Bezeichnung *r.e.* = *recursively enumerable* = *rekursiv-aufzählbar*, in der neueren Literatur auch *c.e.* = *computably enumerable*, erklärt sich aus der folgenden Charakterisierung:

Eine nicht-leere Teilmenge $A \subseteq \mathbb{N}$ ist partiell entscheidbar gdw sie Wertebereich einer rekursiven Funktion ist, also von einer rekursiven Funktion aufgezählt wird:

$$A\ r.e. \iff A = \emptyset \lor A = \{f(x) \mid x \in \mathbb{N}\} \text{ für eine rekursive Funktion } f : \mathbb{N} \to \mathbb{N}.$$

Jede rekursive Relation ist auch r.e.; die Umkehrung gilt, wenn auch das Komplement r.e. ist:

$A \subseteq \mathbb{N}^k$ ist entscheidbar gdw $A$ und $\mathbb{N}^k - A$ sind r.e.

Ist wieder $\varphi_n$ die $n$-te berechenbare Funktion (in einer geeigneten effektiven Aufzählung), so sind die Mengen

$$\begin{array}{ll} \{n \mid \varphi_n(n) \downarrow\} & \text{r.e., aber nicht rekursiv, während} \\ \{n \mid \varphi_n(n) \uparrow\} & \text{nicht einmal r.e. ist.} \end{array}$$

Von besonderer Bedeutung wird später sein, daß für Theorien $\mathsf{T}$ mit einer rekursiven Menge von Axiomen die Folgerungsmenge $\mathsf{C}(\mathsf{T}) = \{\sigma \mid \mathsf{T} \vdash \sigma\}$ r.e. ist. Insbesondere ist also eine rekursiv-axiomatisierbare vollständige Theorie entscheidbar!

# Kapitel 15

# Definierbarkeit berechenbarer Funktionen

## 15.1 Eine endlich-axiomatisierbare Teiltheorie von PA

Die Theorie $\mathsf{PA}$ enthält mit dem Induktionsschema unendlich-viele Axiome, und es läßt sich nachweisen, daß diese nicht durch endlich-viele ersetzbar sind. Trotzdem genügt es, für manche Überlegungen eine geeignete Teiltheorie zugrundezulegen, die nur endlich-viele Axiome benötigt, etwa die folgende Theorie $\mathsf{PA}^-$:

Diese Theorie faßt die wichtigsten algebraischen Eigenschaften der Zahlentheorie in einer Sprache $\mathcal{L}$ mit den Symbolen $0, 1, <, +, \cdot$ zusammen: Die ersten Axiome besagen, daß Addition und Multiplikation assoziativ und kommutativ sind sowie das Distributivgesetz erfüllen, ferner daß 0 und 1 neutrale Elemente für die jeweiligen Operationen und 0 ein Nullteiler ist:

| | | | |
|---|---|---|---|
| A1 | $(x+y)+z = x+(y+z)$ | A6 | $x+0 = x \wedge x \cdot 0 = 0$ |
| A2 | $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ | A7 | $x \cdot 1 = 1$ |
| A3 | $x+y = y+x$ | | |
| A4 | $x \cdot y = y \cdot x$ | | |
| A5 | $x \cdot (y+z) = x \cdot y + x \cdot z$ | | |

Dabei benutzen wir die in der Algebra üblichen Klammeregeln ($\cdot$ bindet stärker als $+$). Für die $<$-Beziehung gelten die Gesetze einer linearen Ordnung, welche mit Addition und Multiplikation verträglich ist:

A8 $\neg\, x < x$

A9 $x < y \wedge y < z \rightarrow x < z$

A10 $x < y \vee x = y \vee y < x$
A11 $x < y \to x + z < y + z$
A12 $0 < z \wedge x < y \to x \cdot z < y \cdot z$

Eine Zahl kann von einen größeren subtrahiert werden:

A13 $x < y \to \exists z\,(x + z = y)$

und schließlich ist 1 ist der Nachfolger von 0 und 0 das kleinste Element (wobei wie üblich $x \leq y :\iff x = y \vee x < y$):

A14 $0 < 1 \wedge \forall x\,(0 < x \to 1 \leq x)$
A15 $\forall x\,(0 \leq x)$

Aus den letzten Axiomen folgt, daß allgemeiner $x+1$ der Nachfolger von $x$ ist und damit die Ordnung *diskret* ist:

$$x < x+1 \wedge \forall y\,(x < y \to x+1 \leq y).$$

**Einige Modelle von $\mathsf{PA}^-$**

1. In der PEANO-Arithmetik $\mathsf{PA}$ kann man die $<$-Beziehung definieren durch $x < y \leftrightarrow \exists z\,(x + z + 1 = y)$ und damit alle Axiome von $\mathsf{PA}^-$ erhalten. Insbesondere ist das Standardmodell $\mathbb{N}$ mit der gewöhnlichen $<$-Beziehung, den üblichen Operationen und den natürlichen Zahlen 0,1 ein Modell von $\mathsf{PA}^-$.

2. Die Menge $\mathbb{Z}[X]$ der Polynome in einer Unbestimmten $X$ und mit ganzzahligen Koeffizienten ist mit den üblichen Operationen ein kommutativer Ring. Man kann diesen Ring ordnen, indem man für ein Polynom $p = a_nX^n + \ldots a_1X + a_0$ mit höchstem Koeffizienten $a_n \neq 0$ setzt:

$$a_nX^n + \ldots a_1X + a_0 > 0 :\iff a_n > 0$$

   und damit Polynome $p, q \in \mathbb{Z}[X]$ ordnet durch $p < q :\iff q - p > 0$.

   Die Teilmenge $\mathbb{Z}[X]^+$ der Polynome $p \in \mathbb{Z}[X]$ mit $p \geq 0$ wird damit zu einem Modell von $\mathsf{PA}^-$, in welchem das Polynom $X$ größer als alle konstanten Polynome und damit “unendlich groß” ist.

In einem Ring liegt bezüglich der Addition eine Gruppe vor, während A13 nur eine eingeschränkte Inversenbildung zuläßt. Ersetzt man in $\mathsf{PA}^-$ die Axiome A13 und A15 durch das Axiom

A16 $\forall x\, \exists z\, (x+z=0)$

so erhält man die algebraische Theorie DOR der **diskret geordneten Ringe**, deren Modelle z. B. die Ringe $\mathbb{Z}$ und $\mathbb{Z}[X]$ sind. Jedes Modell $\mathcal{M}$ von $\mathsf{PA}^-$ kann man zu einem Modell $\mathcal{R}$ der Theorie DOR erweitern (nach demselben Muster, wie man die natürlichen Zahlen zum Ring der ganzen Zahlen erweitert), so daß die nicht-negativen Elemente von $\mathcal{R}$ mit dem ursprünglichen Modell übereinstimmen. Umgekehrt ist auch für jedes Modell $\mathcal{R}$ der Theorie DOR die Einschränkung auf die nicht-negativen Elemente ein Modell von $\mathsf{PA}^-$, so daß man $\mathsf{PA}^-$ als *Theorie (des nicht-negativen Teils) diskret geordneter Ringe bezeichnen* kann.

## 15.2 Arithmetische Formeln

Für Terme $t$ und Formeln $\varphi$ der Sprache $\mathcal{L}$ schreiben wir

$$\begin{array}{lll} \exists x < t\ \varphi & \text{für} & \exists x(x<t \wedge \varphi) \\ \forall x < t\ \varphi & \text{für} & \forall x(x<t \to \varphi) \end{array}$$

und nennen $\exists x < t$ und $\forall x < t$ **beschränkte** Quantoren. Endliche Folgen von Variablen kürzen wir durch $\vec{x}, \vec{y}, \ldots$ ab.

**Definition**

$$\begin{array}{lll} \varphi \text{ ist } \Delta_0\text{-Formel:} & \iff & \varphi \text{ enthält höchstens beschränkte Quantoren,} \\ \varphi \text{ ist } \Sigma_1\text{-Formel:} & \iff & \varphi = \exists\vec{x}\ \psi \quad \text{für eine } \Delta_0\text{-Formel } \psi, \\ \varphi \text{ ist } \Pi_1\text{-Formel:} & \iff & \varphi = \forall\vec{x}\ \psi \quad \text{für eine } \Delta_0\text{-Formel } \psi, \end{array}$$

Dieses ist der Anfang der **Hierarchie der arithmetischen Formeln**; setzt man

$$\Sigma_0 = \Pi_0 = \Delta_0,$$

so kann man fortsetzen:

$$\begin{array}{lll} \varphi \text{ ist } \Sigma_{n+1}\text{-Formel} & \iff & \varphi = \exists\vec{x}\ \psi \quad \text{für eine } \Pi_n\text{-Formel } \psi, \\ \varphi \text{ ist } \Pi_{n+1}\text{-Formel} & \iff & \varphi = \forall\vec{x}\ \psi \quad \text{für eine } \Sigma_n\text{-Formel } \psi. \end{array}$$

Eine $\Sigma_3$-Formel ist also von der Form $\exists\vec{x}\ \forall\vec{y}\ \exists\vec{z}\ \psi$, wobei $\psi$ höchstens beschränkte Quantoren enthält. Das bedeutet also, daß man beschränkte Quantoren nicht mitzählt, mit $\Sigma$ bzw. $\Pi$ anzeigt, daß die Formel mit einer (endlichen) Folge von $\exists$-Quantoren bzw. $\forall$-Quantoren beginnt und der Index die Quantorenblöcke zählt. Es kommt also weniger auf die Anzahl der Quantoren an als die Anzahl der Quantorenwechsel.

Bei dieser Klassifizierung unterscheidet man nicht zwischen logisch äquivalenten Formeln, so daß etwa jede $\Pi_n$-Formel für $n < m$ auch eine $\Sigma_m$- und $\Pi_m$-Formel ist (indem man der Formel einfach zusätzliche Quantoren über nicht vorkommende Variablen voranstellt). Somit kann man dann auch die Formelmengen

$$\Delta_n = \Sigma_n \cap \Pi_n$$

definieren. Es ergibt sich daraus folgendes Bild der **arithmetischen Hierarchie**:

$$\begin{tikzcd}
 & & \Sigma_1 \arrow[dr, dash] & & \Sigma_2 \arrow[dr, dash] & & {} & \\
\Delta_0 \arrow[r, dash] & \Delta_1 \arrow[ur, dash] \arrow[dr, dash] & & \Delta_2 \arrow[ur, dash] \arrow[dr, dash] & & \Delta_3 \arrow[ur, dash] \arrow[dr, dash] & & \cdots\cdots \\
 & & \Pi_1 \arrow[ur, dash] & & \Pi_2 \arrow[ur, dash] & & {} &
\end{tikzcd}$$

Viele nützliche Eigenschaften natürlicher Zahlen lassen sich mittels $\Delta_0$-Formeln ausdrücken, z. B.:

$$x \textit{ ist irreduzibel} \leftrightarrow 1 < x \wedge \forall u < x \, \forall v < x \, \neg(u \cdot v = x).$$

Von besonderer Bedeutung ist, daß $\Delta_0$-Formeln in einer Struktur dasselbe bedeuten wie in allen Enderweiterungen:

## 15.3 Enderweiterungen

Bereits bei der Behandlung von Nonstandard-Modellen von $\mathsf{PA}$ hatten wir bemerkt, daß sich das Standardmodell in jedes andere Modell von $\mathsf{PA}$ einbetten läßt. Diese Eigenschaft gilt auch für die Modelle von $\mathsf{PA}^-$:

### Definition

$\mathcal{L}$ sei eine Sprache, welche ein 2-stelliges Symbol $<$ enthalte, $\mathcal{M}$ und $\mathcal{N}$ seien $\mathcal{L}$-Strukturen mit $\mathcal{M} \subseteq \mathcal{N}$. Dann heißt $\mathcal{N}$ **Enderweiterung** von $\mathcal{M}$ (und entsprechend $\mathcal{M}$ **Anfangssegment** von $\mathcal{M}$) genau dann, wenn die größere Menge $N$ unterhalb eines Elementes von $M$ kein weiteres hinzufügt:

$$\mathcal{M} \subseteq_{end} \mathcal{N} :\Longleftrightarrow \text{ für alle } x \in M, y \in N : (y <^N x \Rightarrow y \in M).$$

Jede natürliche Zahl $n$ wird im Standardmodell, das wir hier auch einfach mit $\mathbb{N}$ bezeichnen, durch den konstanten Term

$$\underline{n} = 1 + \ldots + 1 \quad \text{(n-mal)}$$

dargestellt, wobei $\underline{0}$ die Konstante 0 ist.

### Satz

*Es sei* $\mathcal{M} \models \mathsf{PA}^-$*. Dann wird durch die Abbildung* $n \mapsto \underline{n}^{\mathcal{M}}$ *eine Einbettung des Standardmodells* $\mathbb{N}$ *auf ein Anfangssegment von* $\mathcal{M}$ *definiert.*

*Insbesondere ist (mittels Identifizierung) jedes Modell von* $\mathsf{PA}^-$ *isomorph zu einer Enderweiterung des Standardmodells* $\mathbb{N}$.

*Beweis:* Durch einfache Induktion zeige man für alle natürlichen Zahlen $n, k, l$:

$$\begin{aligned}
n = k + l &\implies \mathsf{PA}^- \vdash \underline{n} = \underline{k} + \underline{l} \\
n = k \cdot l &\implies \mathsf{PA}^- \vdash \underline{n} = \underline{k} \cdot \underline{l} \\
n < k &\implies \mathsf{PA}^- \vdash \underline{n} < \underline{k} \\
\mathsf{PA}^- \vdash \forall x\, (\, x \leq \underline{k} &\rightarrow \quad x = \underline{0} \vee \ldots \vee x = \underline{k})
\end{aligned}$$

Die ersten 3 Aussagen werden wir später für alle rekursiven Funktionen bzw. Relationen verallgemeinern; aus ihnen folgt, daß die Abbildung $n \mapsto \underline{n}^{\mathcal{M}}$ ein Homomorphismus und wegen der letzten Aussage eine Einbettung auf ein Anfangssegment von $\mathcal{M}$ ist. $\square$

## 15.4 Erhaltungseigenschaften unter Enderweiterungen

*$\mathcal{N}, \mathcal{M}$ seien Strukturen der Sprache $\mathcal{L}$ von* $\mathsf{PA}^-$, $\mathcal{N} \subseteq_{end} \mathcal{M}, \vec{a} \in N$. *Dann gilt:*

(i) *jede $\Delta_0$-Formel $\varphi(\vec{v})$ ist absolut:*

$$\mathcal{N} \models \varphi[\vec{a}] \iff \mathcal{M} \models \varphi[\vec{a}],$$

(ii) *jede $\Sigma_1$-Formel $\varphi(\vec{v})$ ist aufwärts-persistent:*

$$\mathcal{N} \models \varphi[\vec{a}] \Longrightarrow \mathcal{M} \models \varphi[\vec{a}],$$

(iii) *jede $\Pi_1$-Formel $\varphi(\vec{v})$ ist abwärts-persistent:*

$$\mathcal{M} \models \varphi[\vec{a}] \Longrightarrow \mathcal{N} \models \varphi[\vec{a}],$$

(iv) *jede $\Delta_1$-Formel $\varphi(\vec{v})$ ist absolut:*

$$\mathcal{N} \models \varphi[\vec{a}] \iff \mathcal{M} \models \varphi[\vec{a}].$$

*Beweis* von (i) durch Induktion über den Formelaufbau von $\varphi(\vec{v})$, wobei nur der Fall eines beschränkten Quantors von Interesse ist. Da aber $\mathcal{M}$ Enderweiterung von $\mathcal{N}$ ist, werden von $M$ unterhalb eines Elementes von $N$ keine neuen Elemente eingefügt, so daß ein beschränkter Quantor in beiden Strukturen dasselbe besagt. □

Wir wollen nun zeigen, daß die rekursiven Relationen mit den Mengen übereinstimmen, die sich durch eine $\Delta_1$-Formeln in den natürlichen Zahlen definieren lassen, und daß der Graph einer rekursiven Funktion durch eine $\Sigma_1$-Formeln in den natürlichen Zahlen definiert werden kann. Mit $\mathbb{N}$ bezeichnen wir im folgenden auch das Standardmodell $(\mathbb{N}, +, \cdot, ', 0, 1)$.

## 15.5 Lemma

*Für jede $\Delta_0$-Formel $\theta(\vec{v})$ ist die Relation*

$$R(\vec{a}) \iff \mathbb{N} \models \theta(\vec{a})$$

*primitiv-rekursiv.*

*Beweis:* Wir zeigen durch Induktion über $\mathrm{lz}(\theta)$, daß die zugehörige charakteristische Funktion

$$c_\theta(\vec{x}) = \begin{cases} 1 & \text{falls } \mathbb{N} \models \theta(\vec{x}) \\ 0 & \text{sonst} \end{cases}$$

primitiv-rekursiv ist: Zunächst sind die Funktionen $x+1, x+y, x \cdot y$ p.r. und damit definiert jeder Term in $\mathbb{N}$ eine primitiv-rekursive Funktion. Da auch die Funktionen $eq(x,y) = \overline{sg}(|x-y|)$ und $sg(y \dot{-} x)$ primitiv-rekursiv sind (und die p.r. Funktionen abgeschlossen sind unter Substitution), gilt die Behauptung für die atomaren Formeln $t = s, t < s$. Für den Fall der aussagenlogischen Operationen benutze man

$$c_{\neg\theta}(\vec{x}) = 1 \dot{-} c_\theta(\vec{x}),\ c_{\theta\wedge\psi}(\vec{x}) = \min(c_\theta(\vec{x}), c_\psi(\vec{x})),\ c_{\theta\vee\psi}(\vec{x}) = \max(c_\theta(\vec{x}), c_\psi(\vec{x})).$$

Ist schließlich $\theta$ eine $\Delta_0$-Formel, $t$ ein Term und $\psi(\vec{x}) = \forall y < t(\vec{x})\ \theta(\vec{x}, y)$, so folgt die Behauptung aus

$$c_\psi(\vec{x}) = eq(t(\vec{x}), (\mu y < t(\vec{x})(c_\theta(\vec{x}, y) = 0)).$$

Ähnlich argumentiert man im Falle der Formel $\exists y < t(\vec{x})\ \theta(\vec{x}, y)$ (oder führt diesen Fall mittels der Negation auf den früheren zurück). $\square$

Die Umkehrung des obigen Lemmas gilt nicht: es gibt primitiv-rekursive Mengen, die durch keine $\Delta_0$-Formel in den natürlichen Zahlen definierbar sind. - Dagegen kann man die folgenden Prädikate und Funktionen als p.r. nachweisen:

$$\begin{array}{rcll} x|y: & \leftrightarrow & \exists z \le y\ (x \cdot z = y) & x \textit{ teilt } y \\ P(x): & \leftrightarrow & \forall u \le x\,(u|x \rightarrow u = x \vee u = 1) & x \textit{ ist Primzahl} \\ p_n & = & (n+1)\textit{te Primzahl}: \quad p_0 = 2, p_1 = 3, p_3 = 5, p_4 = 7, p_5 = 11\ldots & \end{array}$$

Damit können wir endliche Mengen natürlicher Zahlen durch Zahlen codieren, z. B.

$$(n_0, n_1, \ldots, n_k) \quad \textit{durch} \quad p_0^{n_0+1} \cdot p_1^{n_1+1} \cdot \ldots \cdot p_k^{n_k+1}.$$

Definieren wir für $x, y \in \mathbb{N}$

$$\langle x, y \rangle := \frac{(x+y)(x+y+1)}{2} + y,$$

so erhalten wir eine **Paarfunktion** $\mathbb{N}^2 \leftrightarrow \mathbb{N}$, die man durch Iteration erweitern kann, indem man

$$\langle x_1, x_2, \ldots x_k \rangle = \langle x_1, \langle x_2, \ldots x_k \rangle \rangle$$

setzt. Tatsächlich benötigen wir aber zusätzlich eine Funktion, die uniform jede endliche Menge natürlicher Zahlen durch eine einzelne Zahl codiert:

## 15.6 Gödels Lemma

*Es gibt es eine primitiv-rekursive Funktion $\beta$, so daß für jedes k und für jede endliche Folge $x_0, x_1, \dots x_{k-1} \in \mathbb{N}$ gilt: es gibt eine natürliche Zahl c mit*

$$\textit{für alle } i < k: \; \beta(c,i) = x_i.$$

*Tatsächlich gibt es eine* $\Delta_0$-Formel $\theta(x,y,z)$, *so daß*

$$\mathbb{N} \models \forall x,y \, \exists! z \, \theta(x,y,z),$$

*und die Formel $\theta(x,y,z)$ über den natürlichen Zahlen die Funktion $\beta$ definiert.*

*Beweis:* Es sei $x_0, x_1, \dots x_{k-1}$ eine Folge natürlicher Zahlen.
Setze $m := b!$ mit $b := \max(k, x_0, x_1, \dots x_{k-1})$. Dann ist die Folge der Zahlen

$$m+1, 2m+1, \dots k \cdot m + 1$$

paarweise teilerfremd. Nach dem Satz über simultane Kongruenzen (*Chinesischer Restsatz*) gibt es eine natürliche Zahl $a$ mit

$$a \equiv x_i \quad mod((i+1)m+1) \quad \text{für alle } i < k.$$

Nun kann man $\langle a, m \rangle$ als Code der Folge $x_0, x_1, \dots x_{k-1}$ wählen, denn hieraus erhält man für jedes $i < k$ die $x_i$ zurück als Rest der Division von $a$ durch die Zahl $(i+1)m+1$. Bezeichnet $rest(x:y) = z$ den Rest bei der Division von $x$ durch $y$ (falls $y \neq 0$ und $rest(x:0) = 0$), so ist dieses eine p.r. Funktion mit einer $\Delta_0$-Definition. Eine weitere p.r. Funktion erhalten wir durch

$$\alpha(a,m,i) = rest(a : (1+i)m+1).$$

Bezeichnen $p_1, p_2$ die Umkehrfunktionen der obigen Paarfunktion $\langle x, y \rangle$, für die also gilt

$$\langle p_1(x), p_2(x) \rangle = x, \quad \text{wobei } p_1(x), p_2(x) \leq x,$$

so sind auch diese p.r., und wir erhalten schließlich die **Gödelsche $\beta$-Funktion** als

$$\beta(c,i) = \alpha(p_1(c), p_2(c), i).$$

$\square$

Für eine partielle Funktion $f$ bezeichne $\Gamma_f$ den **Graphen** von $f$, also die Relation

$$\Gamma_f(\vec{x}, y) :\Longleftrightarrow f(\vec{x}) = y.$$

## 15.7 Definierbarkeitssatz für rekursive Funktionen

*Eine partielle Funktion $f$ ist berechenbar genau dann, wenn ihr Graph $\Gamma_f$ über den natürlichen Zahlen durch eine $\Sigma_1$-Formel definierbar ist, d. h. wenn es eine $\Sigma_1$-Formel $\theta$ gibt, so daß für alle $\vec{x}, y \in \mathbb{N}$ gilt:*

$$\Gamma_f(\vec{x}, y) \iff \mathbb{N} \models \theta[\vec{x}, y].$$

*Beweis:* Wir zeigen zunächst den einfacheren Teil ($\Leftarrow$): Sei $\theta = \exists \vec{z}\, \psi$ eine $\Sigma_1$-Formel, welche den Graphen $\Gamma_f$ einer partiellen Funktion $f$ definiere, $\psi(\vec{x}, \vec{z}, y)$ also eine $\Delta_0$-Formel. Dann erhalten wir nach Lemma 15.5 eine berechenbare Funktion $g$ durch

$$g(\vec{x}) = \mu z(\mathbb{N} \models \exists y, \vec{z} \leq z\,(\langle y, \vec{z}\rangle = z \wedge \psi(\vec{x}, \vec{z}, y))).$$

Es ist dann $p_1(g(\vec{x}))$ das kleinste $y$ mit $\mathbb{N} \models \theta[\vec{x}, y]$, falls ein solches $y$ existiert, und undefiniert sonst, denn für jedes $\vec{x} \in \mathbb{N}$ gibt es höchstens ein $y$ mit $\mathbb{N} \models \theta[\vec{x}, y]$. Somit ist $p_1(g(\vec{x})) \cong f(\vec{x})$, und damit ist $f$ berechenbar.

Zum Beweis von ($\Rightarrow$) nennen wir partielle Funktionen $f$, deren Graph sich durch eine $\Sigma_1$-Formel über den natürlichen Zahlen definieren lassen, *Funktionen mit einem $\Sigma_1$-Graphen* und zeigen, daß die Menge dieser Funktionen die Anfangsfunktionen enthält und unter Substitution, primitiver Rekursion und unter dem $\mu$-Operator abgeschlossen ist. Dabei beschränken wir uns auf den einzigen Schritt, der etwas schwieriger ist und die GÖDELsche $\beta$-Funktion benutzt: Die Funktion $f$ entstehe aus Funktionen $g, h$ durch primitive Rekursion:

$$\begin{aligned} f(\vec{x}, 0) &\cong g(\vec{x}) \\ f(\vec{x}, y+1) &\cong h(\vec{x}, y, f(\vec{x}, y)) \end{aligned}$$

Dabei können wir annehmen, daß die Graphen von $g$ und $h$ durch $\Sigma_1$-Formeln definierbar sind. Um damit den Graphen von $f$ zu beschreiben, beachten wir, daß sich der Wert von $f(\vec{x}, y)$ berechnet durch die vorangegangenen Werte

$$u_0 = f(\vec{x}, 0), u_1 = f(\vec{x}, 1), \ldots, u_y = f(\vec{x}, y),$$

die wir mittels der $\beta$-Funktion durch eine einzige Zahl $u$ codieren können. Wir müssen nun aber in dieser Folge die Funktion $f$ eliminieren, um eine *explizite* Definition zu erhalten. Der erste Wert $u_0$ ist durch den Graphen von $g$ festgelegt ist und die weiteren $u_{i+1}$ bestimmen sich aus $u_i$ gemäß den Rekursionsbedingungen mittels des Graphen von $h$:

$$\forall i < y \exists r,s \in \mathbb{N}\, [\beta(u,i) = r \wedge \beta(u,i+1) = s \wedge \Gamma_h(\vec{x},i,r,s)]$$

Diese Formel läßt sich (indem man ein genügend großes $w$ wählt) umformen zu

$$\exists w \in \mathbb{N}\, \forall i < y \exists r,s < w\, [\beta(u,i) = r \wedge \beta(u,i+1) = s \wedge \Gamma_h(\vec{x},i,r,s)].$$

Der letzte Wert der obigen Folge, die durch die Zahl $u$ codiert wird, ist der Funktionswert von $f$ an der Stelle $(\vec{x},y)$. Damit erhalten wir als Beschreibung des Graphen von $f$:

$$\begin{aligned}\Gamma_f(\vec{x},y,z) \iff\ & \exists u,v,w \in \mathbb{N}\, (\Gamma_g(\vec{x},v) \wedge \beta(u,0) = v \wedge \beta(u,y) = z \wedge \\ & \forall i < y\, \exists r,s < w\, [\beta(u,i) = r \wedge \beta(u,i+1) = s \wedge \Gamma_h(\vec{x},i,r,s)]), \\ \iff\ & \mathbb{N} \models \Theta[\vec{x},y,z]\end{aligned}$$

für eine $\Sigma_1$-Formel $\Theta$. □

Nennt man eine Menge $A \subseteq \mathbb{N}^k$, die durch eine Formel in $\Gamma$ über den natürlichen Zahlen definierbar ist, ein $\Gamma$-*Menge*, so erhalten wir eine neue Charakterisierung der ersten Stufen der arithmetischen Hierarchie:

## Korollar

(i) *Die $\Sigma_1$-Mengen sind die r.e. Relationen $R \subseteq \mathbb{N}^k$,*

(ii) *Die $\Pi_1$-Mengen sind die Komplemente von r.e. Relationen,*

(iii) *die $\Delta_1$-Mengen sind die die rekursiven Relationen $R \subseteq \mathbb{N}^k$.*

Eine Variante der Definierbarkeitsergebnisse, die sich auf die Beweisbarkeit in formalen Theorien bezieht, ist etwas schwieriger zu behandeln, ermöglicht es aber, Grenzen der Beweisbarkeit aufzuzeigen:

## 15.8 Repräsentierbarkeit

$\mathsf{T}$ sei eine Theorie der zahlentheoretischen Sprache $\mathcal{L}$. Eine (totale) Funktion $f : \mathbb{N}^k \to \mathbb{N}$ ist in $\mathsf{T}$ **repräsentierbar** gdw es eine $\mathcal{L}$-Formel $\theta(x_1,\ldots,x_k,y)$ gibt, so daß für alle $n_1,\ldots,n_k,m \in \mathbb{N}$

(a) $\mathsf{T} \vdash \exists! y \theta(\underline{n_1}, \ldots, \underline{n_k}, y)$ und

(b) $f(n_1, \ldots, n_k) = m \Longrightarrow \mathsf{T} \vdash \theta(\underline{n_1}, \ldots, \underline{n_k}, \underline{m})$.

Ebenso ist eine Menge $S \subseteq \mathbb{N}^k$ in der Theorie $\mathsf{T}$ **repräsentierbar** gdw es eine $\mathcal{L}$-Formel $\theta(x_1, \ldots, x_k)$ gibt, so daß für alle $n_1, \ldots, n_k \in \mathbb{N}$

(c) $(n_1, \ldots, n_k) \in S \Rightarrow \mathsf{T} \vdash \theta(\underline{n_1}, \ldots, \underline{n_k})$ und

(d) $(n_1, \ldots, n_k) \notin S \Rightarrow \mathsf{T} \vdash \neg\theta(\underline{n_1}, \ldots, \underline{n_k})$.

Ist die Funktion $f$ (oder die Menge $S$) durch eine $\Sigma_1$-Formel repräsentierbar, so heißt $f$ (bzw. $S$) **$\Sigma_1$-repräsentierbar**. Die Definition der Repäsentierbarkeit ist so gewählt, daß diese Eigenschaft erhalten bleibt, wenn man die Theorie $\mathsf{T}$ erweitert.

**Satz**

(i) *Jede rekursive Funktion $f : \mathbb{N}^k \to \mathbb{N}$ ist in* $\mathsf{PA}^-$ $\Sigma_1$*-repräsentierbar.*

(ii) *Jede rekursive Menge $S \subseteq \mathbb{N}^k$ ist in* $\mathsf{PA}^-$ $\Sigma_1$*-repräsentierbar.*

*Beweis von* (i): $f : \mathbb{N}^k \to \mathbb{N}$ sei eine rekursive Funktion, ihr Graph $\Gamma_f$ ist also über den natürlichen Zahlen durch eine $\Sigma_1$-Formel $\exists \vec{z}\, \varphi(\vec{x}, y, \vec{z})$ definierbar, wobei $\varphi$ nur beschränkte Quantoren enthält. Da jede Formel der Form $\exists \vec{z}\, \varphi$ in $\mathsf{PA}^-$ äquivalent ist zu $\exists u\, \exists \vec{z}\, (\vec{z} < u \wedge \varphi)$, können wir annehmen, daß $\vec{z}$ nur eine einzelne Variable $z$ ist. Wir bilden nun die $\Delta_0$-Formel $\psi(\vec{x}, y, z)$

$$\varphi(\vec{x}, y, z) \wedge \forall u, v \leq y + z\, (u + v < y + z \to \neg\varphi(\vec{x}, u, v))$$

und weisen nach, daß die $\Sigma_1$-Formel $\exists z\, \psi(\vec{x}, y, z)$ die Funktion $f$ in $\mathsf{T}$ repräsentiert:

Zunächst zeigen wir (b) und nehmen dazu an, daß $f(n_1, \ldots, n_k) = m$ gilt, also $\mathbb{N} \models \exists z\, \varphi(\underline{n_1}, \ldots, \underline{n_k}, \underline{m}, z)$. Die Zahl $m$ ist dabei als Funktionswert eindeutig bestimmt. Wähle $l$ als die kleinste Zahl, so daß $\mathbb{N} \models \varphi(\underline{n_1}, \ldots, \underline{n_k}, \underline{m}, \underline{l})$. Dann gilt offenbar auch $\mathbb{N} \models \psi(\underline{n_1}, \ldots, \underline{n_k}, \underline{m}, \underline{l})$, also auch $\mathbb{N} \models \exists z \psi(\underline{n_1}, \ldots, \underline{n_k}, \underline{m}, \underline{z})$. Als $\Sigma_1$-Satz bleibt dieser Satz bei allen Enderweiterungen des Standardmodells zu einem Modell von $\mathsf{PA}^-$ erhalten, gilt also in allen Modellen von $\mathsf{PA}^-$, und somit gilt $\mathsf{PA}^- \vdash \exists z \psi(\underline{n_1}, \ldots, \underline{n_k}, \underline{m}, \underline{z})$ nach dem Vollständigkeitssatz.

Der Nachweis von (a) benutzt ein ähnliches Argument: Es sei $f(n_1, \ldots, n_k) = m$ und $l$ wieder die kleinste Zahl, so daß $\mathbb{N} \models \psi(\underline{n_1}, \ldots, \underline{n_k}, \underline{m}, \underline{l})$ gilt. Sei $\mathcal{M} \models \mathsf{PA}$. Dann ist zu zeigen, daß $m$ das einzige Element von $M$ ist, welches die Formel $\psi(\underline{n_1}, \ldots, \underline{n_k}, x, \underline{l})$ in $\mathcal{M}$ erfüllt. Daß es überhaupt in $\mathcal{M}$ die Formel erfüllt, liegt an der Absolutheit von $\Delta_0$-Formeln. Sind $a, b \in M$ zwei Elemente, so daß $\mathcal{M} \models$

$\psi(\underline{n_1},\ldots,\underline{n_k},a,b)$, so müssen (nach der Definition von $\psi$) $a,b \leq k+l$ und damit in $\mathbb{N}$ sein. Damit gilt dann wieder wegen der Absolutheit von $\Delta_0$-Formeln auch $\mathbb{N} \models \psi(\underline{n_1},\ldots,\underline{n_k},a,b)$, woraus $\mathbb{N} \models \exists z\, \psi(\underline{n_1},\ldots,\underline{n_k},a,z)$ und damit $a = m$ folgt.

Der Teil (ii) folgt nun hieraus: Ist $S$ rekursiv, so ist die charakteristische Funktion $c_S$ berechenbar, also nach (i) durch eine $\Sigma_1$-Formel $\theta(\vec{x},y)$ in $\mathsf{PA}^-$ repräsentiert. Dann wird $S$ repräsentiert durch die Formel $\theta(\vec{x},1)$, denn es gilt

$$\mathsf{PA}^- \vdash \neg\theta(\vec{x},1) \iff \mathsf{PA}^- \vdash \theta(\vec{x},0).$$

$\square$

# Kapitel 16

# Die Gödelschen Sätze

Die beiden GÖDELschen Unvollständigkeitssätze beruhen auf

1. einer *Codierung* der Formeln der Sprache durch Zahlen (*Gödel-Nummern)*, wodurch syntaktische Eigenschaften von Formeln als zahlentheoretische Aussagen in der Sprache der PEANO-Arithmetik selbst ausgedrückt werden können und

2. einem *Diagonalverfahren*.

## 16.1 Gödel-Nummern

Zunächst ordnen wir jedem Term $t$ der Sprache $\mathcal{L}$ von $\mathsf{PA}^-$ eine Zahl, die GÖDEL-Nummer $\ulcorner t \urcorner$, durch primitive Rekursion wie folgt zu:

$$\begin{aligned}
\ulcorner 0 \urcorner &= 1 \\
\ulcorner 1 \urcorner &= 3 \\
\ulcorner v_i \urcorner &= 3^2 \cdot 5^i \\
\ulcorner s+t \urcorner &= 3^3 \cdot 5^{\ulcorner s \urcorner} \cdot 7^{\ulcorner t \urcorner} \\
\ulcorner s \cdot t \urcorner &= 3^4 \cdot 5^{\ulcorner s \urcorner} \cdot 7^{\ulcorner t \urcorner}
\end{aligned}$$

So ist z. B. $\ulcorner v_3 + 1 \urcorner = 3^3 \cdot 5^{3^2 \cdot 5^3} \cdot 7^3$. Ähnlich kann man nun den Formeln ihre jeweilige GÖDEL-Nummer zuordnen:

$$\begin{aligned}
\ulcorner s = t \urcorner &= 2 \cdot 5^{\ulcorner s \urcorner} \cdot 7^{\ulcorner t \urcorner} \\
\ulcorner s < t \urcorner &= 2 \cdot 3 \cdot 5^{\ulcorner s \urcorner} \cdot 7^{\ulcorner t \urcorner} \\
\ulcorner \neg\varphi \urcorner &= 2 \cdot 3^2 \cdot 5^{\ulcorner \varphi \urcorner} \\
\ulcorner \varphi \vee \psi \urcorner &= 2 \cdot 3^3 \cdot 5^{\ulcorner \varphi \urcorner} \cdot 7^{\ulcorner \psi \urcorner} \\
\ulcorner \exists v_i \varphi \urcorner &= 2 \cdot 3^4 \cdot 5^i \cdot 7^{\ulcorner \varphi \urcorner}
\end{aligned}$$

Natürlich gibt es auch viele andere Möglichkeiten einer Codierung; wichtig ist hier nur, daß es sich um eine berechenbare injektive Abbildung aller $\mathcal{L}$-Formeln[1] auf eine berechenbare Teilmenge $Fml \subseteq \mathbb{N}$ handelt und daß die die Umkehrabbildung (definiert auf der Menge $Fml$) ist ebenfalls berechenbar ist (so daß wir aus der GÖDEL-Nummer $\ulcorner \varphi \urcorner$ die Formel $\varphi$ wieder ausrechnen können).

Im folgenden benötigen wir die **Diagonal-Funktion** $d$, definiert durch

$$d(n) = \begin{cases} \ulcorner \forall y\,(y = \underline{n} \to \sigma(y)) \urcorner, & \text{falls } n = \ulcorner \sigma(v_0) \urcorner \text{ für eine } \mathcal{L}\text{-Formel } \sigma(v_0) \\ 0 & \text{sonst.} \end{cases}$$

## 16.2 Diagonalisierungslemma

*$\mathsf{T}$ sei eine $\mathcal{L}$-Theorie, $\theta(v_0)$ eine $\mathcal{L}$-Formel mit der einen freien Variablen $v_0$. Falls die Diagonalfunktion $d$ in $\mathsf{T}$ repräsentierbar ist, so gibt es einen $\mathcal{L}$-Satz $G$ mit der Eigenschaft*

$$\mathsf{T} \vdash G \leftrightarrow \theta(\underline{\ulcorner G \urcorner}).$$

*Ist $d$ durch eine $\Sigma_1$-Formel in $\mathsf{T}$ repräsentierbar und $\theta(v_0)$ eine $\Pi_1$-Formel, so kann $G$ als $\Pi_1$-Satz gewählt werden.*

*Beweis:* Ist $d$ in $\mathsf{T}$ durch die Formel $\delta(x,y)$ repräsentiert, so definiere

$$\psi(v_0) := \forall y\,(\delta(v_0, y) \to \theta(y)),$$

[1] Berechenbare Funktionen haben wir hier nur als zahlentheoretische Funktionen eingeführt, es ist jedoch nahe liegend, diesen Begriff auch auf andere Bereiche (wie endliche Folgen von Symbolen) auszudehnen.

und setze $n := \ulcorner \psi(v_0) \urcorner$. (Achtung: $n$ ist tatsächlich eine Zahl, in welcher natürlich die Variable $v_0$ nicht vorkommt; dagegen geht die GÖDEL-Nummer von $v_0$ in die Berechnung von $n$ ein!) Für $G$ wählt man nun den Satz

$$G := \forall y\,(y = \underline{n} \to \psi(y)).$$

Dann hat also $G$ die GÖDEL-Nummer $d(\ulcorner \psi(v_0) \urcorner)$. Ist $\delta(x,y)$ eine $\Sigma_1$-Formel und $\theta(v)$ eine $\Pi_1$-Formel, so sind $\psi$ und $G$ (äquivalent zu) $\Pi_1$-Formeln. Somit bleibt zu zeigen, daß $\mathsf{T} \vdash G \leftrightarrow \theta(\underline{\ulcorner G \urcorner})$: Es gilt offenbar

$$\begin{array}{lll} & \mathsf{T} \vdash G \leftrightarrow \psi(\underline{n}), & \text{d. h. nach Def. von } \psi \\ (1) & \mathsf{T} \vdash G \leftrightarrow \forall y(\delta(\underline{n},y) \to \theta(y)). & \end{array}$$

Daß $d$ in $\mathsf{T}$ durch $\delta(x,y)$ repräsentiert wird, besagt aber

$$d(n) = \ulcorner G \urcorner \Rightarrow \mathsf{T} \vdash \delta(\underline{n}, \underline{\ulcorner G \urcorner}) \quad \text{und} \quad \mathsf{T} \vdash \exists! y\, \delta(\underline{n}, y).$$

Somit gilt insbesondere

$$\begin{array}{lll} (2) & \mathsf{T} \vdash \forall y\,(\delta(\underline{n},y) \leftrightarrow y = \underline{\ulcorner G \urcorner}) & \text{und zusammen mit (1) folgt daraus} \\ & \mathsf{T} \vdash G \leftrightarrow \forall y\,(y = \underline{\ulcorner G \urcorner}) \to \theta(y)), & \text{also} \\ & \mathsf{T} \vdash G \leftrightarrow \theta(\underline{\ulcorner G \urcorner}). & \end{array}$$

□

## 16.3 Satz von Gödel-Rosser

$\mathsf{T}$ *sei eine rekursive Menge von (Gödel-Nummern von) $\mathcal{L}$-Sätzen, so daß gilt:*

(i) $\mathsf{T}$ *enthält keinen Widerspruch, d. h. es gibt keinen $\mathcal{L}$-Satz $\sigma$ mit $\ulcorner \sigma \urcorner \in \mathsf{T}$ und zugleich $\ulcorner \neg\sigma \urcorner \in \mathsf{T}$,*

(ii) $\mathsf{T}$ *enthält alle $\Sigma_1$ und $\Pi_1$-Sätze, die in $\mathsf{PA}^-$ gelten:*

$$\{\ulcorner \sigma \urcorner \mid \sigma \in \Sigma_1 \cup \Pi_1, \sigma \text{ Satz}, \mathsf{PA}^- \vdash \sigma\} \subseteq \mathsf{T}.$$

*Dann ist $\mathsf{T}$ $\Pi_1$-unvollständig in folgendem Sinne: es gibt einen $\Pi_1$-Satz $\tau$ mit*

$$\ulcorner \tau \urcorner \notin \mathsf{T} \text{ und } \ulcorner \neg\tau \urcorner \notin \mathsf{T}.$$

*Beweis:* Da $\mathsf{T}$ rekursiv ist, gibt es eine $\Sigma_1$-Formel $\theta(v)$, so daß für alle $n \in \mathbb{N}$

$$\begin{aligned} n \in \mathsf{T} &\implies \mathsf{PA}^- \vdash \theta(\underline{n}), \\ n \notin \mathsf{T} &\implies \mathsf{PA}^- \vdash \neg\theta(\underline{n}). \end{aligned}$$

Nach dem Diagonalisierungslemma 16.2 gibt es einen $\Pi_1$-Satz $\tau$ mit

$$\mathsf{PA}^- \vdash \tau \leftrightarrow \neg\theta(\underline{\ulcorner\tau\urcorner}).$$

Mit (ii) folgt daraus

$$\ulcorner\tau\urcorner \in \mathsf{T} \Rightarrow \mathsf{PA}^- \vdash \theta(\underline{\ulcorner\tau\urcorner}) \Rightarrow \mathsf{PA}^- \vdash \neg\tau \Rightarrow \ulcorner\neg\tau\urcorner \in \mathsf{T}$$

und ebenso mit (i)

$$\ulcorner\neg\tau\urcorner \in \mathsf{T} \Rightarrow \ulcorner\tau\urcorner \notin \mathsf{T} \Rightarrow \mathsf{PA}^- \vdash \neg\theta(\underline{\ulcorner\tau\urcorner}) \Rightarrow \mathsf{PA}^- \vdash \tau \Rightarrow \ulcorner\tau\urcorner \in \mathsf{T},$$

so daß weder $\ulcorner\tau\urcorner$ noch $\ulcorner\neg\tau\urcorner$ in $\mathsf{T}$ sein können. □

Eine Theorie $\mathsf{T}$ heißt **rekursiv-axiomatisierbar** gdw sie ein Axiomensystem $\Sigma$ besitzt, für welches die Menge der GÖDELnummern $\{\ulcorner\sigma\urcorner \mid \sigma \in \Sigma\}$ rekursiv ist.

## 16.4 Erster Gödelscher Unvollständigkeitssatz

$\mathsf{T}$ *sei eine konsistente*[2] *und rekursiv-axiomatisierbare Theorie der Sprache* $\mathcal{L}$*, welche die Theorie* $\mathsf{PA}^-$ *erweitert. Dann ist* $\mathsf{T}$ $\Pi_1$*-unvollständig, d. h. es gibt einen* $\Pi_1$*-Satz* $\tau$ *mit*

$$\mathsf{T} \nvdash \tau \quad \textit{und} \quad \mathsf{T} \nvdash \neg\tau.$$

*Beweis:* Wir nehmen an, daß $\mathsf{T}$ sowohl konsistent und rekursiv-axiomatisierbar als auch $\Pi_1$-vollständig ist und leiten daraus einen Widerspruch ab: Setze

$$S := \{\ulcorner\sigma\urcorner \mid \sigma \in \Sigma_1 \cup \Pi_1, \sigma \text{ Satz }, \mathsf{T} \vdash \sigma\}.$$

Dann ist $S$ rekursiv (denn um nachzuprüfen, ob ein $\Sigma_1$ oder $\Pi_1$-Satz $\sigma$ in $S$ ist, braucht man nur die rekursive Aufzählung der in $\mathsf{T}$ beweisbaren Sätze der Form

[2] Die ursprüngliche Fassung des GÖDELschen Satzes hatte die stärkere Voraussetzung der $\omega$-Konsistenz von $\mathsf{T}$ (d. h. für alle $\mathcal{L}$-Formeln $\theta(v)$ mit $\mathsf{T} \vdash \theta(\underline{n})$ für alle $n \in \mathbb{N}$ ist $\mathsf{T}[\forall v\, \theta(v)]$ konsistent). ROSSER hat 1936 die obige stärkere Version bewiesen.

$\sigma$ bzw. $\neg\sigma$ zu durchlaufen, wegen der $\Pi_1$-Vollständigkeit muß einer der beiden nach endlich-vielen Schritten auftauchen). Somit läßt sich der Satz von GÖDEL-ROSSER anwenden, wonach $S$ $\Pi_1$-unvollständig ist, ein Widerspruch! □

Jede konsistente Erweiterung $\mathsf{T}$ von $\mathsf{PA}^-$ hat nach dem Satz von LINDENBAUM 4.9.1 eine vollständige und konsistente Erweiterung, von denen also keine rekursiv (d. h. entscheidbar) sein kann. Aus dem obigen Satz kann man übrigens folgern, daß jede rekursive konsistente Erweiterung von $\mathsf{PA}^-$ sogar $2^{\aleph_0}$-viele Erweiterungen und ebenso viele abzählbare Modelle besitzt, die nicht elementar-äquivalent sind (s. R. KAYE: *Models of Peano Arithmetic,* pp. 39ff).

Mathematisch interessantere Sätze, die in der PEANO-Arithmetik unabhängig sind, erhält man durch

## 16.5 Kombinatorische Prinzipien

In seiner einfachsten Form besagt das

**Schubfachprinzip:**

(1) *Verteilt man n Elemente in m < n-viele Schubladen, so muß eine der Schubladen mindestens 2 Elemente enthalten.*

Die Verallgemeinerung auf unendliche Mengen lautet:

(2) *Ist eine unendliche Menge zerlegt in endlich-viele Mengen, so muß wenigstens eine dieser Mengen unendlich sein:*

$$A = A_0 \dot\cup \ldots \dot\cup A_k \textit{ unendlich } \rightarrow \exists i \leq k\ (A_i \textit{ unendlich})\ .$$

Zur weiteren Verallgemeinerung setzen wir

$[A]^n := \{x \subseteq A \mid |x| = n\}$ Menge der $n$-elementigen Teilmengen von $A$.

Eine *Zerlegung* der $n$-elementigen Teilmengen von $A$ in $k$ Teilmengen kann man auch durch eine Funktion

$$f : [A]^n \rightarrow k$$

wiedergeben ($A_i = \{x \in A \mid f(x) = i\}$ sind dann die Zerlegungsmengen), und ein solches $f$ nennt man etwas anschaulicher auch eine **Färbung** (von $[A]^n$) mit $k$ Farben. Eine Teilmenge $X \subseteq A$ mit $[X]^n \subseteq A_i$ für ein $i < k$ (deren $n$-Tupel also einfarbig sind) heißt **homogen** für die Zerlegung $f$. Der

**Satz von Ramsey**[3]

$$f : [\mathbb{N}]^n \to k \;\to \exists X \subseteq \mathbb{N}\,(X \; unendlich \wedge \forall u,v \in [X]^n f(u) = f(v)).$$

besagt also, daß jede Färbung der $n$-elementigen Teilmengen von natürlichen Zahlen mit endlich-vielen Farben eine unendliche homogene Teilmenge besitzt. Dieser Satz ist in der Mengenlehre mit einer schwachen Form des Auswahlaxioms beweisbar (s. etwa D. MARKER: *Model Theory: An Introduction,* § 5.1).

In diesem Zusammenhang benutzt man folgende Notationen:

**Definition**

Eine (endliche) Menge $H$ von Ordinalzahlen heißt

$$\textbf{relativ groß} \leftrightarrow card(H) \geq min(H).$$

Sind $n,k \in \mathbb{N}, \alpha,\gamma \in On$, so bedeute

$$\gamma \longrightarrow (\alpha)^n_k \quad bzw. \quad \gamma \xrightarrow{min} (\alpha)^n_k$$

daß es für jede Färbung $f$ der $n$-elementigen Teilmengen von $\gamma$ mit $k$ Farben eine (relativ große) Teilmenge $H \subseteq \gamma$ vom Ordnungstyp $\alpha$ gibt, die homogen für $f$ ist.

Der obige Satz von RAMSEY besagt also: $\forall n,k\; \omega \longrightarrow (\omega)^n_k$. Hieraus folgert man den *finitären Satz von* RAMSEY:

$$\forall m,n,k \exists r\; r \longrightarrow (m)^n_k.$$

Dieser Satz über natürliche Zahlen läßt sich in der Sprache der PEANO-Arithmetik ausdrücken und auch beweisen; eine (anscheinend) leichte Verstärkung aber nicht mehr: PARIS-HARRINGTON[4] haben gezeigt:

**Satz**

Die Aussage

$$\forall m,n,k \exists r\; r \xrightarrow{min} (m)^n_k$$

ist in der Mengenlehre beweisbar, nicht aber in PA (falls PA konsistent ist). Die Existenz dieser Zahlen $r$ kann man in PA nicht mehr allgemein zeigen, tatsächlich führen sie zu sehr großen Zahlen: Ist

$$\sigma(n) = das\; kleinste\; r\; mit\; r \xrightarrow{min} (n+1)^n_n,$$

[3] F. P. RAMSEY: *On a problem of formal logic.* Proc. London Math. Soc. (2) 30 (1929), 264-286

[4] *A mathematical Incompleteness in Peano Arithmetic.* In: Handbook of Mathematical Logic, ed. Barwise, Elsevier 1982

so übertrifft die Funktion $\sigma$ schließlich jede rekursive Funktion $f$ (d. h. es zu jeder rekursiven Funktion $f$ gibt es ein $p$ mit $f(n) < \sigma(n)$ für alle $n \geq p$).[5]

## 16.6 Zweiter Gödelscher Unvollständigkeitssatz

Mit Hilfe einer GÖDEL-Nummerierung kann man syntaktische Eigenschaften von Formeln in zahlentheoretische Eigenschaften übersetzen. Da auch Beweise endliche Folgen von Formeln sind, kann man auch diese durch natürliche Zahlen codieren und $\mathcal{L}$-Formeln $Bew_{\mathsf{T}}, bewb_{\mathsf{T}}$ finden, so daß

(i) $Bew_{\mathsf{T}}(n,m)$ ausdrückt: "n ist die Nummer eines Beweises einer Formel $\varphi$ mit der GÖDELnummer $m$" und

(ii) $bewb_{\mathsf{T}}(m) :\leftrightarrow \exists n Bew_{\mathsf{T}}(n,m)$ ausdrückt: "m ist die GÖDELnummer einer Formel $\varphi$, welche in $\mathsf{T}$ beweisbar ist."

Ist die Theorie $\mathsf{T}$ rekursiv-axiomatisierbar, so ist das Beweisprädikat rekursiv und das Beweisbarkeitsprädikat r.e. Ferner kann man dann für eine geeignete $\Sigma_1$-Formel $\theta = bewb_{\mathsf{T}}$ die folgenden **Ableitbarkeits-Bedingungen** nachweisen:

$$\begin{aligned} \mathsf{T} \vdash \sigma \quad &\Longrightarrow \quad \mathsf{T} \vdash \theta(\ulcorner\sigma\urcorner) \\ \mathsf{T} \vdash \theta(\ulcorner\sigma \to \tau\urcorner) \quad &\to \quad (\theta(\ulcorner\sigma\urcorner) \to \theta(\ulcorner\tau\urcorner)) \\ \mathsf{T} \vdash \theta(\ulcorner\sigma\urcorner) \quad &\to \quad \theta(\ulcorner(\theta(\ulcorner\sigma\urcorner)\urcorner) \end{aligned}$$

(Dabei sind $\sigma, \tau$ beliebige Sätze der Sprache $\mathcal{L}$, und die GÖDEL-Nummern $n$ sollten in diesen Formeln jeweils durch die entsprechenden Terme $\underline{n}$ ersetzt sein. Zum Nachweis benötigt man allerdings eine stärkere Theorie als $\mathsf{PA}^-$, etwa $\mathsf{PA}$.) Unter Benutzung eines derartigen Beweisbegriffes kann man auch die *Konsistenz* einer Theorie formalisieren, etwa

$$Con_{\mathsf{T}} :\leftrightarrow \neg\theta(\ulcorner 0 = 1\urcorner).$$

Damit kann man dann zeigen:

[5] J. KETONEN - R. SOLOVAY: *Rapidly growing Ramsey functions.* Annals of mathematics (2) 113 (1981), 267-314. Einen guten Überblick über dieses Thema findet man auch bei
C. SMORYNSKI: *Some rapidly growing functions.* The Mathematical Intelligencer 2,3 (1979), 149-154

## 2. Gödelscher Unvollständigkeitssatz

$\mathsf{T}$ *sei eine Theorie der Sprache* $\mathcal{L}$, *für welche ein formales Beweisprädikat* $\theta = \theta_{\mathsf{T}}$ *mit den obigen Bedingungen existiere (z. B.* $\mathsf{T} = \mathsf{PA}$). *Dann gilt:*

$$\mathsf{T} \textit{ ist konsistent} \Longrightarrow \mathsf{T} \nvdash \mathit{Con}_{\mathsf{T}}.$$

*Beweis:* (Skizze) Nach dem Diagonalisierungslemma existiert ein Satz $\tau$, so daß:

$$(*) \qquad \mathsf{T} \vdash \tau \leftrightarrow \neg\theta(\underline{\ulcorner\tau\urcorner}).$$

Falls $\mathsf{T} \vdash \tau$, so folgt aus der 1. Ableitbarkeitsbedingung: $\mathsf{T} \vdash \theta(\underline{\ulcorner\tau\urcorner})$, also nach (*) $\mathsf{T} \vdash \neg\tau$, d. h. $\mathsf{T}$ ist inkonsistent. Damit haben wir:

$$(**) \qquad \mathsf{T} \textit{ konsistent} \Longrightarrow \mathsf{T} \nvdash \tau.$$

Man zeige, daß sich dieser Beweis formalisieren und mit Hilfe des durch $\theta$ in $\mathsf{T}$ formalisierten Beweisprädikates darstellen läßt, also

$$\mathsf{T} \vdash \mathit{Con}_{\mathsf{T}} \to \neg\theta_{\mathsf{T}}(\underline{\ulcorner\tau\urcorner}).$$

Sei nun $\mathsf{T}$ konsistent, also nach (**) $\mathsf{T} \nvdash \tau$. Dann ist wegen (*) auch $\mathsf{T} \nvdash \neg\theta(\underline{\ulcorner\tau\urcorner})$ und somit auch $\mathsf{T} \nvdash \mathit{Con}_{\mathsf{T}}$. □

Dieser Satz besagt also, daß man selbst mit den Mitteln der PEANO-Arithmetik keinen Beweis der Konsistenz dieser Theorie finden kann (ein Beweis der Konsistenz einer Theorie $\mathsf{T}$ würde überdies nur überzeugend sein, wenn er unter schwächeren Voraussetzungen möglich wäre, als sie Theorie $\mathsf{T}$ zur Verfügung stellt), und diese Aussage gilt auch für alle weiteren formalen Theorien, in welcher sich die PEANO-Arithmetik interpretieren läßt (wie z. B. die übliche axiomatische Mengenlehre).

## 16.7 Wahrheit ist nicht arithmetisch definierbar

Während sich die üblichen syntaktischen Begriffe, die zu einer formalen Sprache $\mathcal{L}$ gebildet werden, in der Sprache der Zahlentheorie definieren lassen und ihre wesentlichen Eigenschaften zumindest in der Theorie PA beweisen lassen, ist dies für den semantischen Wahrheitsbegriff nicht möglich:

## Satz von Tarski

*Es sei* $\mathcal{M} \models \mathsf{PA}^-$*. Dann gibt es keine* $\mathcal{L}$*-Formel* $\theta(v)$*, so daß für alle natürlichen Zahlen* $n$

$$\mathcal{M} \models \theta(\underline{n}) \iff n = \ulcorner\sigma\urcorner \text{ für einen } \mathcal{L}\text{-Satz } \sigma \textit{ mit } \mathcal{M} \models \sigma.$$

*Beweis:* Wenn es eine solche Wahrheitsdefinition $\theta$ gäbe, so könnte man nach dem Diagonalisierungslemma 16.2 einen Satz $G$ finden, so daß

$$\begin{array}{rcl} \mathsf{PA}^- \vdash G & \leftrightarrow & \neg\theta(\underline{G}), \quad \textit{insbesondere} \\ \mathcal{M} \models G & \iff & \mathcal{M} \models \neg\theta(\underline{\ulcorner G\urcorner}). \end{array}$$

Andererseits müßte für die Wahrheitsdefinition gelten:

$$\mathcal{M} \models G \iff \mathcal{M} \models \theta(\underline{\ulcorner G\urcorner}),$$

was aber zum Widerspruch führt! □

## 16.8 Unentscheidbarkeit

Wie im vorigen Abschnitt werden wir gelegentlich eine formale Theorie $\mathsf{T}$ mit

$$\ulcorner\mathsf{T}\urcorner = \{\ulcorner\sigma\urcorner \mid \sigma \ \textit{ Satz mit } \mathsf{T} \vdash \sigma\},$$

der Menge der GÖDEL-Nummern der in $\mathsf{T}$ beweisbaren Sätze identifizieren. So heißt eine Theorie $\mathsf{T}$ **entscheidbar** gdw $\ulcorner\mathsf{T}\urcorner$ entscheidbar (d. h. rekursiv) ist; ist $\mathsf{T}$ nicht entscheidbar, so heißt $\mathsf{T}$ **unentscheidbar**. Für eine entscheidbare Theorie $\mathsf{T}$ gibt es also ein effektives Verfahren, welches nach endlich-vielen Schritten feststellt, ob ein vorgegebener Satz (in der Sprache der Theorie) in $\mathsf{T}$ beweisbar ist oder nicht.

Indem man die Diagonalfunktion $d$ benutzt, kann man nach der Methode des Satzes von GÖDEL-ROSSER 16.3 zeigen:

## Satz

*Ist* $\mathsf{T}$ *eine konsistente Theorie der Sprache* $\mathcal{L}$*, so sind nicht die Diagonalfunktion* $d$ *und zugleich die Menge* $\ulcorner\mathsf{T}\urcorner$ *in* $\mathsf{T}$ *repräsentierbar.*

*Beweis:* Wir nehmen an, daß sowohl $d$ durch eine Formel $\delta(x,y)$ wie auch die Menge $\ulcorner\mathsf{T}\urcorner$ durch eine Formel $\varphi_{\mathsf{T}}(v_0)$ in $\mathsf{T}$ repräsentiert werden und wählen $\theta = \neg\varphi_{\mathsf{T}}$. Aus dem Diagonalisierungslemma 16.2 erhalten wir dann die Existenz einer Formel $G$ mit

$$(*) \qquad \mathsf{T} \vdash G \leftrightarrow \neg\varphi_{\mathsf{T}}(\underline{\ulcorner G\urcorner}).$$

1. Fall: $\mathsf{T} \vdash G$, also $\ulcorner G\urcorner \in \ulcorner\mathsf{T}\urcorner$. Dann gilt wegen der Repräsentierbarkeit: $\mathsf{T} \vdash \varphi_{\mathsf{T}}(\underline{\ulcorner G\urcorner})$ und mit (*) auch $\mathsf{T} \vdash \neg G$, d. h. $\mathsf{T}$ wäre widerspruchsvoll.

2. Fall: $\mathsf{T} \nvdash G$, also $\ulcorner G\urcorner \notin \ulcorner\mathsf{T}\urcorner$. Dann gilt wiederum wegen der Repräsentierbarkeit: $\mathsf{T} \vdash \neg\varphi_{\mathsf{T}}(\underline{\ulcorner G\urcorner})$, mit (*) also $\mathsf{T} \vdash G$, Widerspruch!

$\square$

## Folgerungen und Bemerkungen

(i) *Ist* $\mathsf{T}$ *eine konsistente Theorie der Sprache* $\mathfrak{L}$, *welche* $\mathsf{PA}^-$ *erweitert, so ist* $\mathsf{T}$ *unentscheidbar.*
(Denn dann ist jede rekursive Funktion, also auch $d$, in $\mathsf{T}$ repräsentierbar.) Da dieses für die endlich-axiomatisierbare Theorie $\mathsf{PA}^-$ gilt, so erhalten wir:

(ii) *Die Prädikatenlogik, d. h.* $\{\ulcorner\sigma\urcorner \mid \sigma \text{ Satz mit } \vdash \sigma\}$, *ist unentscheidbar.*
(CHURCH 1936) Tatsächlich gilt bereits:

(iii) *Die Prädikatenlogik mit einem 2-stelligen Relationszeichen ist unentscheidbar (ebenso die PL mit mindestens zwei 1-stelligen oder mindestens einem 2-stelligen Funktionszeichen).* (TRACHTENBROT 1953)

Denn es gibt eine endlich-axiomatisierbare Teiltheorie der Mengenlehre (in der Sprache mit der $\in$-Relation), in welcher sich alle rekursiven Funktionen repräsentieren lassen. Dagegen ist

(iv) *die monadische Prädikatenlogik (mit nur 1-stelligen Relationszeichen außer der Gleichheit) entscheidbar.* (BEHMANN 1922)

(v) *Die Theorie der Gruppen ist unentscheidbar* (MALCEV 1961), *die Theorie der* ABEL*schen Gruppen dagegen ist entscheidbar* (W. SMIELEW 1949).

Ein Beispiel eines Entscheidungsverfahrens für eine algebraische Theorie behandeln wir im nächsten Teil.

## Das Entscheidungsproblem für die Prädikatenlogik

Ohne Beweis (s. hierzu etwa BÖRGER-GRÄDEL-GUREVICH 1997) führen wir die folgenden Ergebnisse an:

- **entscheidbar** (bezüglich Allgemeingültigkeit) sind Formeln der Prädikatenlogik mit folgenden Präfixtypen:

$$\forall\ldots\forall\exists\ldots\exists$$
$$\forall\ldots\forall\exists\forall\ldots\forall$$
$$\forall\ldots\forall\exists\exists\forall\ldots\forall \qquad \text{(ohne =-Zeichen!)}$$

- **unentscheidbar** (bezüglich Allgemeingültigkeit) sind bereits Formeln der Prädikatenlogik mit folgendem Präfix (selbst ohne =-Zeichen):

$$\exists\forall\exists$$

- **Reduktionstypen** (d. h. Allgemeingültigkeit ist auf Formeln dieses Typs zurückzuführen) sind die folgenden Formelklassen:

$$\exists\exists\exists\forall\ldots\forall$$
$$\exists\exists\forall\ldots\forall\exists$$
$$\exists\exists\forall\exists\ldots\exists$$
$$\forall\exists\forall\exists\ldots\exists$$

# Kapitel 17

# Literatur

## Geschichte

BOCHENSKI, J.M.: *Formale Logik.* Alber, Freiburg 1966
DUMMETT, M.: *Frege.* Harvard Univ. Press 1991
DRUCKER, TH. (ed.): *Perspectives on the history of mathematical logic.* Birkhäuser 1991
KUTSCHERA, F. V.: *Gottlob Frege. Eine Einführung in sein Werk.* de Gruyter 1989
HASENJAEGER, G.: *Grundbegriffe und Probleme der math. Logik.* Alber, Freiburg 1962
KNEALE, W. und M.: *The development of logic.* Oxford Univ. Press 1962
SCHOLZ, H.: *Abriß der Logik.* Alber, Freiburg 1996

## Mathematische Logik

ASSER, G.: *Einführung in die mathematische Logik I-III.* Teubner 1981
EBBINGHAUS-FLUM-THOMAS: *Einführung in die mathematische Logik.* Spectrum Akademischer Verlag 1996
HILBERT-BERNAYS: *Grundlagen der Mathematik I,II.* Springer 1934/39, Neuauflage: 1968/70
HILBERT-ACKERMANN: *Grundzüge der mathematischen Logik.* Springer (mehrere Auflagen)
MATES, B.: *Elementare Logik.* Vandenhoeck & Ruprecht 1978
MENDELSON, E.: *Introduction to Math. Logic.* Chapman & Hall 1997

NOVIKOV, P.S.: *Grundzüge der Mathematischen Logik.* vieweg 1973
PRESTEL, A.: *Einführung in die Mathematische Logik und Modelltheorie.* vieweg 1986
RAUTENBERG, W.: *Einführung in die Mathematische Logik.* vieweg 2002
RASIOWA-SIKORSKI: *The Mathematics of Metamathematics.* Polish Scientific Publ. 1968
SHOENFIELD, J.R.: *Mathematical Logic.* Addison-Wesley 1967 TOURLAKIS, G.: *Lectures in Logic and Set Theory.* Cambridge 2003
TUSCHIK-WOLTER: *Mathematische Logik - kurzgefaßt.* Spectrum Akademischer Verlag 2002

## Mengenlehre

CANTOR, G.: *Gesammelte Abhandlungen.* Springer 1980
DEISER, O.: *Einführung in die Mengenlehre.* Springer 2004
EBBINGHAUS, D.: *Einführung in die Mengenlehre.* Spectrum Akademischer Verlag 2003
FELGNER, U. (Herausgeber): *Mengenlehre.* Wiss. Buchgesellschaft 1979
FRAENKEL, A.- BAR HILLEL, Y. - LÉVY, A.: *Foundations of Set Theory.* Amsterdam 1973
FRIEDRICHSDORF, U.-PRESTEL, A.: *Mengenlehre für den Mathematiker*. vieweg 1985
JECH, TH.: *Set Theory.* Springer 2003
KANAMORI, A.: *The mathematical development of set theory from Cantor to Cohen* Bull. Symb. Logic 2,1, 1-71 (1996)
LÉVY, A.: Basic *Set Theory*. Springer 1979

## Modelltheorie

CHANG-KEISLER: *Model Theory.* Elsevier 1990
HODGES, W.: *Model Theory.* Cambridge Univ. Press 1993
KAY, R.: *Models of Peano Arithmetic.* Oxford University Press 1991
MARCJA-TOFFALORI: *A Guide to Classical and Modern Model Theory.* Kluwer 2003
MARKER, D.: *Model Theory. An Introduction* Springer 2002

## Berechenbarkeit, Unentscheidbarkeit, Unvollständigkeit

BÖRGER, E.: *Berechenbarkeit, Komplexität, Logik.* vieweg 1992
BÖGER-GRÄDEL-GUREVICH: *The classical decision problem.* Springer 1997
CUTLAND, N.: *Computability* Cambridge University Press 1980
KAYE, R.: *Models of Peano Arithmetic* Clarendon Press 1991

## Informatik, autom. Beweisen, Varia

BARWISE, J. (ed.): *Handbook of Mathematical Logic.* Elsevier 1982
GABBAY-GUENTHNER (ed.): *Handbook of Philosophical Logic.* Kluwer 2003
FITTING, M.: *First-order logic and automated theorem proving.* Springer 1990
HEINEMANN-WEIHRAUCH: *Logik für Informatiker* Teubner 1991
HOFSTADTER, D.R.: *Gödel, Escher, Bach.*
Basic Book 1979, deutsch bei Klett-Cotta und dtv
SCHÖNING, U.: *Logik für Informatiker* B.I. Verlag 1987

**Bibliographie**

G.H. MÜLLER und V. LENSKI (ed.): *Ω-Bibliography of Mathematical Logic.*
Vol. I-VI. Springer 1987, fortgesetzt im internet unter:
http://www-logic.uni-kl.de/BIBL/index.html
Weitere links auch über http://www.math.uni-heidelberg.de/logic/wwwlinks.html

**Skripten im Internet** über die links
http://www.math.uni-heidelberg.de/logic/skripten.html
http://www.uni-bonn.de/logic/world.html
Biographien: http://www-groups.dcs.st-and.ac.uk/ history/BiogIndex.html

**Zeitschriften über Mathematische Logik**

Annals of Pure and Applied Logic
Archive of Mathematical Logic
Fundamenta Mathematicae
Journal of Symbolic Logic, Bulletin of Symbolic Logic
Mathematical Logic Quarterly (früher: Zeitschrift f. math. Logik und Grundl. der Math.)
Notre Dame Journal of Formal Logic